

IVF remains in legal limbo

The British government, first in the field with a coherent policy on research in innovatory human biology, is lapsing into sloth by its seeming willingness to let legislation be delayed.

WHAT (if anything) should be done to regulate research and medical practice in human embryology? For most of this decade, that has been a lively and contentious public issue. It will remain so for many years to come, especially as late entrants to the controversy declare their interests. (The Vatican, which issued its instruction on *in vitro* fertilization (IVF) only a few weeks ago, is a special case; its comment on current disputes may have been slow in coming, but its general position has been well-known.) In the circumstances, it may be inevitable that the pattern of regulation is muddled and confusing. In most places, most of the United States, for example, *laissez faire* seems to be the guiding principle. Elsewhere, as in Australia, there are islands of zealotry (the state of Victoria) in a sea of relative indifference. In West Germany, it seems that the cards are being dealt for what will be a battle between the federal justice ministry and professional interests, while in Britain, where the government three years ago welcomed the report of its Warnock committee, promised legislation is still not forthcoming.

The British problem is especially instructive because the issues to be decided have been drawn with admirable clarity. The Warnock committee decided that IVF techniques are a welcome means of treating infertility which nevertheless prompts legal questions about parentage (where eggs or sperm are derived from sources other than the nominal parents), but that public anxiety about the uses that might be made in laboratories of artificially fertilized human embryos requires that research projects in this field should be regulated, preferably intelligently, and that breaches of the regulations should be criminal offences. Both the government and the scientific community in Britain accepted these conclusions; the government promised legislation within a year. Now, three years later, the oddly named Voluntary Licensing Authority (for human IVF and embryology) is complaining (see p.92) that both its own work and research in human embryology are being hampered by the continuing lack of legislation.

Legislation

The circumstances are curious, to say the least. The Voluntary Licensing Authority is the creation of two nominally autonomous organizations, the Medical Research Council and the Royal College of Obstetricians and Gynaecologists; the licences it has issued to about 30 centres have no statutory force. Even so, there is no reason to suppose that centres practising IVF have chosen not to announce their existence. Helpfully, the "authority" also now provides a list of research projects at the licensed centres. The vast majority are directed towards the improvement of the efficiency of IVF techniques (still only 11 per cent per menstrual cycle); the two exceptions (out of 32) are angled at the early detection of chromosome abnormalities, which is part of the same theme. But experience has also shown that some centres active in the field decline to make spare embryos available to others because of the legal uncertainties that persist, which is "an unfortunate waste of unique and precious material". That is another good reason why the uncertainty should be removed.

But how? The British government's position, outlined in a

discussion paper last December, is that it wishes to act deliberately and comprehensively, even if that implies delay. The betting is that a bill will not appear until October 1988, if then. Meanwhile, the danger persists that public policy will be hijacked by some group of private parliamentarians. (Next week, we shall be publishing a characteristically impassioned statement to just such a group by Dr Erwin Chargaff, the molecular biologist.) Part of the trouble is that the government has been unlucky in its short-stay health ministers; Mr Barney Hayhoe, appointed only a few months after Warnock, was conscientiously at odds with the government's intentions, his successor, Ms Edwina Currie, seems more interested in helping people to stop smoking cigarettes and to take up jogging. But the lawyers are also letting their search for perfection cloud their judgement. IVF and the other innovations in human embryology on which decisions are required do not constitute an integral field of public policy in which model laws once written will stand for the rest of time. Rather, they are extensions of present practice that provoke awkward questions in unrelated fields. (Are IVF children legitimate for the purposes of inheritance? Do their nominal parents have a duty to tell them if one gamete or another is derived from some other source? What kinds of research projects are permissible now? And when more is known of the process of development?)

Frameworks

Pragmatic as it is, the British government probably does not differ much from others in its readiness to let these questions lie for the time being, knowing as it must that they cannot be debated without raising all kinds of other awkward questions, the rights and wrongs of abortion law, for example. But there is an urgent need to regularize present practice in IVF (for which the Warnock proposals are sufficient) and to allow generally acceptable research to proceed within a framework that both researchers and the world at large can understand, but which can be changed with circumstances. That is one task.

Another might be to ban the use of the word "pre-embryo", used by the voluntary authority as a synonym for a fertilized human ovum not yet implanted in a uterus. Put simply, this usage is a cop-out, a way of pretending that the public conflict about IVF and other innovations in human embryology can be made to go away by means of an appropriate nomenclature. The fact is that a fertilized human egg is as much deserving of being called an embryo as is a fertilized frog's egg. The essence of the controversy over the new human embryology centres rests on the question when, in the course of development, an embryo commands the legal respect to which free-living people are entitled. The issue turns on the necessity of implantation for development, on analogies (necessarily less persuasive) with the randomness of what happens in real-life procreation and on arguments about the reality of the soul (which to many is a figment of the human imagination). Even those who share the British self-styled voluntary authority's eagerness that IVF should be more widely and efficiently practised will acknowledge that, on the issue of nomenclature, the Vatican is philosophically the more consistent. □



Retreat on principles

The British government's abandonment of its search for a nuclear dump is an awkward legacy.

THE British government's decision (see p. 93) to abandon the search for a disposal site for low-level radioactive waste is a bad business. Whether or not the decision, or its timing, has been prompted by electoral considerations, and in particular by fears that the loyalty of voters in constituencies bordering on the four candidate sites would be unduly strained in the impending general election, giving up the search will hang a millstone around the necks of future governments. And while the immediate effect of the decision will be further to cramp the nuclear energy industry in Britain, the decision will set an awkward precedent for all future public works that give offence to those directly affected by them.

At present, the only licensed nuclear dump in Britain is at a place called Drigg, on the coast of Cumbria in north-west England, a few miles from Britain's larger reprocessing plant, which has been steadily accumulating low-level radioactive wastes since the 1940s. For more than a decade, first the nuclear industry and, more recently, the waste disposal authority called NIREX have been searching for a new site. As the identity of successive candidate sites has become publicly known, there has been a local outcry and the authorities have looked elsewhere. The latest strategy, that of investigating four sites simultaneously, was evidently designed to suggest that misfortune would be equitably dispersed; the site eventually chosen would not merely have been technically the most suitable but also the one best calculated to minimize the inconvenience and risk to people living nearby. The snag, for the government, has been the cost (£15 million will now be written off) and that four local communities have simultaneously been up in arms.

For the rest of Britain, the immediate snag is that there will now be no disposal sites for low-level wastes. Yet at no point in the recent arguments about the safety of these disposal sites have the objectors been able to establish that dumping slightly contaminated rubbish from nuclear plants and laboratories would hazard the safety and health of those living nearby. The Secretary of State for the Environment, Mr Nicholas Ridley, now says that it will be "almost" as economical to dispose of the same rubbish in the still non-existent repository for intermediate waste for which NIREX is also searching, but other governments than the present one will have to pay the bills.

Rightly or wrongly, the government has decided that Britain must continue to generate electricity from uranium; how can it consistently hold to such a policy if it concedes that local objections to unwelcome consequences should be paramount? It is as if the government, having decided that a country such as Britain needs an efficient road transport network, should nevertheless give such weight to the objections of those who prefer not to live near motorways that it agrees that all new roads should be tunnels. The underlying principle, now sacrificed, is that policies for industrial, economic or social development which are democratically agreed should not be impeded by sectional interests, which is not to deny the right of sectional interests to make their objections known at public inquiries and, in some cases, to be compensated for their loss of amenity.

The issue of compensation has become conspicuous in the long wrangle over the waste disposal sites because Mr William Waldegrave, one of the most able of the government's younger ministers who has special responsibility for environmental issues, appears at one stage to have suggested that those not wishing to live near a nuclear dump should be enabled to move elsewhere "at no cost to themselves". This is tantamount to an even more radical concession of the principle that the awkward consequences of desired development are inescapable. The environmentalists are forever echoing Dr Barry Commoner's dictum that "there is no such thing as a free lunch", but the

Waldegrave doctrine, taken literally, is a way of pretending that there is — that the unwelcome consequences of desired developments can be hidden. In reality, developments whose unwelcome consequences are so seriously regarded that people will move house to avoid them are, by that test almost certainly misconceived. Generally applied, the doctrine is also probably a recipe for national bankruptcy. □

Back to Reykjavik?

The US proposal to cut strategic warheads by a half seems warlike to the Soviet Union.

IN any bilateral negotiation, the more successful partner is usually the one who can most vividly appreciate the other's position — who can look from 'over the hill', as strategists were fond of saying in the late 1960s. Last week, the United States bounced back at its bilateral negotiations with the Soviet Union at Geneva with a proposal that each side should cut by a half the number of nuclear warheads deployed by its strategic missiles, and earned for its pains a sour denunciation by the Soviet news agency TASS. There are many in the United States who will be genuinely puzzled that, after a decade's tacit compliance with SALT II (which froze strategic missile numbers), an offer to cut numbers of strategic warheads by a half could give offence. What follows is therefore diffidently offered primarily as an explanation of that phenomenon.

Historically, the argument goes back to 1972, when the United States and the Soviet Union signed the Anti-Ballistic Missile (ABM) Treaty. The two sides then agreed that their security could only be undermined by the further development of effective ABM systems. In practice, the Soviet government of the time was busily protecting Moscow with a missile system to do just that, with the result that the raw ABM treaty allows each side to have two ground-based point-defence systems, one for its national capital and one for its missile launchers. But addenda to the treaty, partly inspired by the disappointments of the US development of the SPRINT fast take-off ABM rocket, cut the number of defensible Soviet sites from two to one and eschewed US anti-missile defences of any kind.

It is easily forgotten how much the world has changed since then. In 1972, in the United States, it was not possible, as it is now, to walk into a retail store and come out with a box containing as much computer power as there had been in the whole world a decade earlier, in 1962. Undoubtedly, as the Pentagon is forever saying, the Soviet Union has upgraded the ABM defence of Moscow since 1972, but not at zero cost, part of which may be that there are few computer stores in the Moscow suburbs. Present alarms about the trading and technological relationship between the United States and Japan should not blind US taxpayers to the improvement of their skills, relative to those of the Soviet Union, in a mere decade and a half.

Soviet people are more conscious of this difference than are those who live in the United States. American sceptics of the Strategic Defense Initiative (SDI), convinced that it will be ineffectual, should wonder why their Soviet colleagues denounce it as immoral, a much more loaded word. Their answer must be that Soviet people worry that even relative failure will yield some advantage to the United States and better early warning certainly, perhaps even radically improved techniques of terminal defence and certainly a ruinously expensive competition in the militarization of space. It must be concluded that, in Soviet eyes, it cannot be preferable to become economically bankrupt in isolation than to be assured of destruction mutually. That is merely one reason why SDI will be as much a stumbling block at Geneva as it was on last November's Sunday afternoon at Reykjavik. And, given Mr Reagan's attachment to the project, that is why the Soviet Union would prefer a deal on intermediate missiles now, waiting for another administration to strike the more important bargain on strategic arms. □

World Bank policy to add ecology to economics

- Environmental protection to gain priority
- Past errors will not be repeated

Washington

THE World Bank has finally conceded to its critics and promised to give higher priority to environmental protection in its lending policy. "Sound ecology is good economics", said the World Bank president, Barber Conable, in a speech announcing the establishment of a top-level environment department employing 60 people within the bank, alongside four regional technical departments and a series of conservation programmes.

The bank, with annual loans of \$12,000 million, is the world's largest source of finance for development projects in the poorer nations. Bank policy has always been that its "decisions to lend must be based exclusively on economic considerations". But pressure to change has grown after some of its large projects had unforeseen effects on the environment.

Most criticized of all was the massive Polonoreste project in Brazil, which provided loans to pave a highway through the tropical forest into the province of Rondonia. The project was intended to open up the forest to thousands of poor farmers. But according to José Lutzenberger, an agronomist based in Porta Allegre who has become the scheme's most noted critic inside Brazil, "their farming methods were unsuitable and the soil was poor. After a few years, many had to sell their smallholdings and move on, even poorer than when they started. Behind them came speculators and big timber companies." Their activities have triggered worldwide concern for the fate of the Amazon's rain forests. Conable himself now concedes that "the bank misread the human, institutional and physical realities of the jungle". Part of the loan was interrupted because Brazil failed to set aside promised reserves for indigenous peoples and for wildlife.

Many other bank projects have been attacked for failing to take their social and environmental consequences into account. In India massive planting of eucalyptus trees has often turned out to be unsuited to local needs, while a \$840-million power station project at Singrauli produced conditions described in US Congress testimony as resembling the "lower circle of Dante's Inferno". And in 1985, the bank funded a cattle development project in Botswana, despite the failure of similar projects, that helped turn grassland into desert.

According to Conable, past errors are

not to be repeated. "If the World Bank has been part of the problem in the past, I intend to make it a leader for finding solutions for the future", he said. A five-year project on environmental threats to 30 of the most vulnerable developing nations will be carried out to prepare a "natural resources balance sheet", so that resources such as topsoil and grass cover, water and "drainage and human skills and traditional lifestyles" can be entered into economic assessment.

At the same time, the bank intends to design a continent-wide initiative against desertification in Africa, to double annual lending to forestry projects, to participate in global programmes to conserve tropical forests, and to explore the possibility of a cooperative project to clean up the Mediterranean.

On a smaller scale, the bank has recently begun to back reserves in Brazil where natural forest products, such as Brazil nuts and rubber, can be sustainably harvested from the forest. Some half a million people in the Amazon basin live by extracting native rubber, but their needs have been largely ignored in development plans.

Environmental groups have warmly welcomed the bank's new policies but find it hard to believe that they are really being given all that they had been demanding for years. One cause of suspicion is that the changes are part of a much wider reorganization of the bank which Conable is having difficulties in carrying through effectively. He intends to reduce staff, and to eliminate levels of bureaucracy to produce a more flexible institution.

Bruce Rich, senior attorney at the Environmental Defense Fund, an organization that has persistently criticized bank policies, questions whether the new department will be created by "shifting dead wood, rather than hiring people trained in natural sciences, ecology and anthropology". If so, he says, "it would be better to have no new department but hire 50 scientists and call them economists". A deeper worry is that the policy changes may be being made primarily to pacify US environmental groups, many of which have their offices within a mile of the bank's Washington headquarters, and will have little effect on loan conditions. Rich insists that environmental groups in the borrower countries must be involved in project planning.

Alun Anderson

Applications for new superconductors are on the horizon

Washington

PRACTICAL uses of the new superconducting ceramics may be much closer than even the most enthusiastic physicists had guessed. Scientists at the IBM Research Division in Yorktown Heights announced on 11 May that they had succeeded in growing a single superconducting crystal, in the form of a thin film, and that it can carry a hundred times more current than the sintered powders made in laboratories previously. The measured current capacity, greater than $100,000 \text{ A cm}^{-2}$ at the temperature of liquid nitrogen, is enough for many large-scale applications, including power transmission.

With higher critical temperatures, greater resistance to magnetic fields, and now the prospect of higher currents, the new materials seem to have the old liquid-helium-cooled superconductors beaten on all counts. Many problems remain to be solved in making usable cables from the brittle ceramics, but materials scientists differ only in their expectation of how long this will take. Dr Praveen Chaudhari, IBM's vice-president of science, says that they chose to make thin films chiefly because they were familiar with the technique and knew how to make high-quality crystals. His research group has ideas for increasing critical currents still further, and for growing bulk crystals of the same standard.

The IBM scientists also cooled their sample to liquid-helium temperature, and measured a current of $2 \times 10^6 \text{ A cm}^{-2}$; this fell to $1 \times 10^6 \text{ A}$ at a magnetic field of 1 Tesla, then fell off more slowly at higher fields. The magnetic properties of the new superconductors have not been studied in detail, but these first measurements suggest that there is no fundamental obstacle to building high-field magnets.

The currents demonstrated by IBM are comparable to what is now attained in niobium-titanium magnets cooled by liquid helium, so this new discovery will profoundly affect the argument over whether to build the Superconducting Super Collider (SSC). High-energy physicists who support the SSC are anxious to get started, and point to the decades of work that were needed to construct the first generation of superconducting magnets, used in the Fermilab Tevatron. But with so many laboratories and so much commercial interest focused on superconductors, development is proceeding rapidly and unpredictably, and it will be hard to persuade Congress to spend billions of dollars on what many now think as outmoded technology. David Lindley

EPA clears the way for release of nitrogen-fixing microbe

Washington

LAST week, the US Environmental Protection Agency (EPA) cleared the way for the environmental release of a bacterium with an enhanced capability to fix nitrogen. The EPA proposes to permit BioTechnica International, a biotechnology company with US headquarters in Cambridge, Massachusetts, to field-test its recombinant version of *Rhizobium mililoti* on a five-acre plot of alfalfa in Pepin County, Wisconsin. The proposal is original in that it represents the first bacteria to be reviewed under regulations first drafted to govern toxic substances.

BioTechnica's recombinant strain of *Rhizobium mililoti* contains plasmids bearing multiple copies of the gene for the production of the enzyme nitrogenase, a trait that amplifies its ability to fix nitrogen. In greenhouse tests on alfalfa, increases in yield of 15–17 per cent are reported.

Farmers have used commercial preparations of the wild-type bacteria for decades to coat seeds of legumes before sowing to boost crop yields. BioTechnica hopes its recombinant bacteria will make significant inroads into this seed-coat market, estimated to be \$400 million a year for alfalfa and soybean alone.

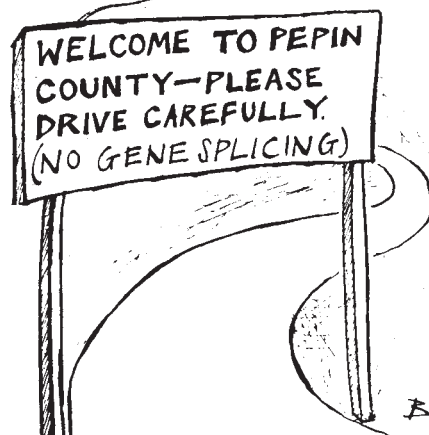
Two other gene-spliced bacteria — a *Pseudomonas* that thwarts ice-nucleation on strawberry plants and another that discourages frost on potato leaves — have already received the go-ahead from EPA to be tested in the field. Advanced Genetic Sciences (AGS) of Oakland, California, sprayed a field of strawberry plants with its "ice-minus" bacteria three weeks ago (see *Nature* 326, 819; 1987), and two weeks ago the University of California (UC) at Berkeley planted seeds treated with its strain of *Pseudomonas* in a potato patch in northern California.

But the BioTechnica bacterium is being put through a different set of regulatory paces than either of the Californian ones, a procedure that disturbs environmental groups and those who would like to see a coherent, unified set of regulations to address specifically the products of biotechnology. The AGS and UC Berkeley bacteria were approved by EPA under the Federal Insecticide, Fungicide and Rodenticide Act (FIFRA), on the grounds that ice on plants is a "pest" and therefore bacteria that combat such ice-formation are "pesticides". According to a spokesperson for EPA, the fortified nitrogen-fixing bacteria from BioTechnica "doesn't kill anything", so it fitted best under the aegis of the Toxic Substances Control Act (TSCA). Both acts are used by EPA to monitor gene-altered microbes under the

Coordinated Framework for Biotechnology announced in June 1986.

Critics of this two-pronged EPA review process claim that reviewing recombinant bacteria under TSCA is inappropriate, because, unlike chemical substances, bacteria have the ability to reproduce and spread.

Under TSCA, a company is required to file a pre-manufacture notification with EPA before making or importing a new chemical substance that is not on EPA's master list. EPA has 90 days to respond to this notification and perform a risk/benefit analysis to determine if the substance



should be regulated. If a response is not forthcoming at the end of the 90 days, the company has tacit approval to make, use and sell the substance without further EPA oversight. Alternatively, EPA can extend the 90-day response period up to an additional 90 days if more information is needed before a decision about the substance can be reached.

EPA has extended its review of the BioTechnica bacteria by 60 days to finalize the details of the field test, and allow for a brief public comment period. The risk/benefit analysis uncovered the possibility that the enhanced nitrogen-fixing bacteria might also accelerate the growth of sweet clover, and cause it to build up coumarin, a chemical that is converted to a substance that impairs blood-clotting in animals when combined with mouldy hay. To prevent this possible side-effect, all sweet clover will be removed from the field-test plot. EPA expects to approve the field test by the end of July.

Local opposition to the field test is mobilized, however, and BioTechnica is likely to face a long haul through public forums and courthouses. Residents are circulating a petition to gather support for a one-year moratorium on the test, during which a state review and full hearings in the county would take place. Carol Ezzell

Supernova 1987a close to peak brightness

Washington

SUPERNOVA 1987a, now with a visual magnitude of about 3.1, may be close to the peak of its optical brightness. Recent observations show that the rate of brightening has slowed down, and at some wavelengths the expanding supernova shell may have begun to dim. But in ultraviolet radiation the brightness is still increasing, and new spectral features are beginning to emerge.

George Sonneborn of Goddard Space Flight Center and Robert Kirshner of Harvard University have been monitoring data from the International Ultraviolet Explorer (IUE) satellite since the supernova was first spotted. For several weeks there has been a steady increase in ultraviolet brightness of about 4 per cent a day, and in the past few days this has been accompanied by the appearance of rather broad lines in the ultraviolet spectrum. With the optical observations, this suggests that the supernova shell is becoming transparent, so that higher energy radiation from within can be seen directly.

These new developments are more or less as expected. Everyone now agrees that the blue supergiant Sk-69 202 really was the progenitor (IUE spectra indicate that only stars 2 and 3 of the original triplet remain), and unpublished calculations by Hillebrandt *et al.*, by Arnett and by Woosley *et al.*, show that the explosion of such a star can produce a 'dim' Type II supernova. In addition, enough ^{60}Ni can be synthesized for radioactive decay to ^{60}Co and then ^{60}Fe to be the cause of the now observed ultraviolet emission.

As the shell continues to expand and rarefy, astronomers may eventually be able to see characteristic gamma rays from the radioactivity within. But the half-life for ^{60}Co to decay to stable ^{60}Fe is about 80 days, so unless the shell becomes transparent quickly, the source of gamma radiation may have expired before it becomes visible. This is one of the ways in which the evolution of the light curve can reveal the internal structure and dynamics of the supernova shell.

An unexplained aspect of the ultraviolet observations is the presence of rather broad spectral features at a number of wavelengths. Sonneborn believes they are due to absorption in the still optically thick shell, and are not emission lines from the interior, but as higher energy photons emerge from the increasingly transparent supernova remnant it should be possible to identify these features more convincingly. David Lindley

Cable and Wireless gains an uneasy foothold in Japan at last

Tokyo

BRITAIN'S Cable and Wireless, with the backing of the British and US governments, has gained a foothold in Japan in its battle to break the monopoly in trans-Pacific telecommunications enjoyed by Kokusai Denshin Denwa (KDD) and American Telephone and Telegraph (AT&T). Japan's Ministry of Post and Telecommunications has agreed in principle to let the British company and its US and Japanese partners lay a trans-Pacific cable. But the catch is that the Cable and Wireless consortium must join hands with an all-Japanese consortium that has totally incompatible plans.

Following the recent uproar in Britain and the United States over Japanese attempts to squeeze Cable and Wireless and US companies out of Japan's international telecommunications market (see *Nature* 326, 319; 1987), the ministry has made some concessions: foreign companies will be allowed to be on the board of directors and an equal share to Japanese partners in a new international telecommunications company to compete with KDD, and the new company can lay a trans-Pacific optical fibre cable, if all partners agree. But therein lies the rub. The ministry continues to insist that International Digital Communications Planning Inc. (IDC) established by the Cable and Wireless group must merge with the all-

Japanese consortium International Telecom Japan Inc. (ITJ) before a licence application to begin service will be considered. A meeting between executives of IDC and ITJ on 1 May failed to agree on a merger plan put forward by Japan's Federation of Economic Organizations (Keidanren) and only served to emphasize the difference between the groups.

Whereas IDC proposes to lay its own cable between Japan and Alaska, ITJ intends to lease circuits on Intelstat satellites and a trans-Pacific cable (TPC-3) to be laid by KDD and AT&T in 1988. IDC will agree to a merger only if ITJ accepts the Japan-Alaska cable plan, and ITJ says the cable should not be a prerequisite for establishing the new company. The two consortia also failed to resolve the question of equity shares. ITJ calls for equal shares for all eight Japanese and foreign partners; IDC maintains it is necessary to form a "core group" of two of three companies holding most shares effectively to manage the firm.

Behind these disagreements lies a much bigger battle between the world's telecommunications giants. The Japan-Alaska cable is part of Cable and Wireless's plan to build a "global digital highway" of optical fibre cables linking Europe, America and the Far East to compete with networks laid by KDD, AT&T and British Telecom.

On 22 April, shortly after Japan's Ministry of Post and Telecommunications indicated conditional approval for IDC's cable, the US Federal Communications Commission granted permission for Pacific Telecom Cable (a joint venture of Cable and Wireless and Pacific Telecom of the United States) to lay the trans-Pacific cable from Alaska. Two days later, KDD and AT&T agreed on a joint survey for laying yet another trans-Pacific cable (TPC-4), even though KDD has been maintaining all along that there is sufficient demand to justify IDC's cable.

TPC-4 is part of a more grandiose plan to link Japan, Australia, New Zealand and the United States via a gigantic loop of optical fibre cables stretching around the Pacific rim that will be laid by KDD, AT&T and telecom companies from Australia and New Zealand. KDD's president, at a press conference last Friday, even invited Cable and Wireless to participate in the laying of TPC-4. But although not without precedent — Cable and Wireless is cooperating with KDD and AT&T in laying a Hong Kong-Japan optical fibre cable — Cable and Wireless seem likely to reject the offer as it has already turned down an option on TPC-3.

David Swinbanks

Nuclear safety lesson

London

BRITAIN'S Nuclear Installations Inspectorate has not been able adequately to study the lessons of the Chernobyl accident because of staff shortages, according to its chief inspector, Eddy Ryder. He told the House of Commons Select Committee on Energy last week that nuclear safety inspectors cannot keep up with international developments because of lack of staff.

The inspectorate has 15 per cent less than its full complement of 120 inspectors, and finds it difficult to attract more staff because of its salary scale. Its limited resources or concentrated on the safety of operating nuclear reactors. K.J.

£50m AIDS fund aim

London

THE fight against AIDS is to be boosted by a new government-backed charity which aims to raise £5 million over two years. Robert Maxwell, chairman of Mirror Group newspapers, will be chief fundraiser and a national AIDS trust, chaired by Sir Austin Bide, president of Glaxo Holdings, will spend the money.

Norman Fowler, Secretary of State for Social Services, stressed that the move is not a substitute for government action, and a coordinated response from both the voluntary sector and the government is essential. S.J.H.

Shuttle widow sues

Washington

THE widow of pilot Michael J. Smith, killed in the explosion of the space shuttle Challenger, has filed a \$1,500 million lawsuit against Morton Thiokol Inc., makers of the defective rocket booster; the US government; and Laurence Mulloy, head of NASA's rocket booster programme at the time of the accident.

The families of four of the seven astronauts killed in the accident have accepted compensation from NASA and Thiokol. One other family is suing Thiokol. The new lawsuit accuses Thiokol and NASA of "reckless disregard for human life". A.A.

Blood products lab

London

A NEW £55-million plasma-processing laboratory which opened last week will see England and Wales self-sufficient in Factor VIII, the product used to treat haemophiliacs, by 1989. A new high-temperature process will inactivate hepatitis non-A non-B and AIDS viruses.

For the first time in the United Kingdom, this drive to produce Factor VIII is expected to lead to surplus production of other plasma products which could be sold. As the Blood Products Laboratory obtains all its plasma from voluntary donations, it stresses that it will not be exploited for commercial purposes. S.J.H.

UCSF space problem

San Francisco

THE University of California at San Francisco (UCSF) may have plenty of staff and money but space is proving harder to find.

Surrounded by a community that is intolerant of expansion, UCSF agreed in 1976 to limit expansion at its urban main campus and decentralize. But its latest plans, to convert office buildings to laboratories for the School of Pharmacy at Laurel Heights, about two miles from the campus, have run into opposition.

Despite a favourable environmental impact report, a community group filed suit to stop the conversion. The group picked out molecular parasitologist Nina Agabian, who planned to move from the Naval Bioscience Laboratory in Oakland. Rumours, strongly denied by the university, circulated that she is researching germ warfare for the Navy. Last September, a superior court judge upheld the environmental impact report. Now, UCSF is awaiting the outcome of an appeal by the neighbourhood group. If the court rules in favour, it will start the move next month.

Marcia Barinaga

Firm resistance to splitting of university research and teaching

London

THE British government's proposals for reorganizing higher education would lead to a system with a limited number of elite universities plus a large number of institutions that will be little more than teaching factories, the new chairman of the Committee of Vice-Chancellors and Principals, Sir Mark Richmond, has warned.

The government last week issued consultative papers calling for comments on a controversial new university funding approach based on contracts, with the abolition of the block grants system run by the University Grants Committee (UGC) and its replacement by a new University

Funding Council (UFC) closely tied to industry. A report recommending the reorganization of the university system for sciences has also been released, with what Richmond termed "considerable implications for the way the university system will be run".

The Earth Sciences Review committee report, published last week by the University Grants Committee (see *Nature* 326, 813; 1987), recommends three levels of institutions — a top group of 10 or 12 universities equipped for international research and teaching, a middle group with research at a lower level without sophisticated equipment, and a bottom group offering only low-level teaching. Although the report of the committee, chaired by Professor Ron Oxburgh, was intended to apply only to geology, its covering letter from the UGC notes that this reorganization has "implications for the future reorganization of other subjects, particularly chemistry and physics". The UGC is asking for reaction from universities, pointing out that implementation would "require close cooperation".

Although some scientists would welcome the concentration of top research work in a few institutions, there are concerns about separating the sciences. "Can you be a top-class research institution if you have only biology, for instance, but not chemistry?" asked Richmond. "I

think you've got to have both."

The system of contracting for funding with the new UFC, set out in the consultative paper, is seen as giving the government increasing power over the direction of universities. The report includes no mechanism for universities to advise the government on their needs, the outgoing chairman of the vice chancellors' committee Mr Maurice Shock said. But it does spell out that the government will offer guidance on "the nation's needs with regards to the size and broad balance of the university system".

Contracts for funding could take one of three forms: a single, comprehensive contract covering all of an institution's activities; a subdivision into separate faculties or departments or even separate courses; or a mesh of these two approaches. "The ferocity of the response to this contracting approach will depend on which option is chosen," Richmond said. "But the touchstone will be value for money. The pressure will be there to do things as cheaply as possible." The intention is to encourage universities to attract contracts from private sources.

The vice chancellors' committee will discuss its reaction to the consultative papers later this month. "I hope these really are consultative documents and that the government will listen to reasoned response," Shock said. But the documents make it clear that legislation to implement the new UFC will be introduced soon, and comments are being sought for "the framing of detailed legislative provisions."

Kathy Johnston

Embryo guidelines opposed by clinics

London

NEW guidelines forbidding donation of eggs by relatives of infertile patients, and the transfer of more than four embryos at a time, were issued last week by the British body that licenses centres carrying out *in vitro* fertilization and embryo research. The new rulings put the Voluntary Licensing Authority (VLA) in conflict with one London clinic, the Wellington Humana Hospital, which last week revealed that two patients who had received eggs from their sisters had recently given birth to twins. The VLA's ruling is in line with its principle that the anonymity of the donor is paramount in the interests of the child.

The licensing authority will meet representatives from the 28 clinics working on *in vitro* fertilization in September to receive assurances that the guidelines will be observed. But without legislation, the only action the VLA can take against clinics that refuse to obey the guidelines is to withdraw its approval; it cannot prevent them from operating.

Announcing the new guidelines, VLA chairman Dame Mary Donaldson stressed the urgent need for legislation and for the establishment of a statutory body with a qualified inspectorate, but held out little hope for progress in the near future.

The VLA was set up two years ago by the Royal College of Obstetricians and Gynaecologists and the Medical Research Council following the completion of the Warnock Report on human fertilization and embryology, which recommended a statutory licensing authority.

Although accepting that some delay was desirable before legislation, the VLA is finding it increasingly difficult to retain sufficient control over the rapidly growing field.

Simon Hadlington

Marc Aaronson (1940–1987)

MARC Aaronson died in an accident at the Mayall 4-m telescope of the National Optical Astronomy Observatories, United States, on 30 April. He received his first degree from Caltech in 1972 and his doctorate from Harvard in 1977, where he worked with several astronomers, including John Danziger, Jay Frogel and Eric Persson. He took part in the effort to understand the infrared energy distribution of galaxies in terms of their stellar populations, and wrote a thesis on the infrared properties of spiral galaxies, showing the increasing contribution of young stars to the integrated optical emission of later type galaxies.

In 1977 he went to the University of Arizona, Tucson, to work. He made several important discoveries, the first of which was the usefulness of the infrared Tully–Fisher relation in measuring redshift-independent relative distances of galaxies. He led a collaboration with Greg Bothun, John Huchra, Paul Schechter, Robert Schommer, Woodruff Sullivan, Brent Tully, myself and others. With these more precise distances to galaxies he

recognized the signal of the Virgo-centric flow within the Local Supercluster, that is, the deceleration of the Hubble flow locally by the mass centred on the Virgo cluster.

In 1978, Aaronson proposed that we follow up the suggestion that carbon stars were responsible for the peculiar optical and infrared colours of some globular clusters in the magellanic clouds. As a result of frequent trips to Cerro Tololo Interamerican Observatory we were able to show that carbon stars were a characteristic of intermediate age stellar populations, as predicted by some theoretical models of stellar evolution.

Aaronson's third important discovery surprised everyone. He measured the velocity dispersion of carbon stars in dwarf spheroidal galaxies, by means of high resolution spectroscopy at the Multiple Mirror Telescope. He found a velocity dispersion significantly greater than the expectations from globular clusters, indicating that even the tiniest dwarf galaxies have dark halos. These results favour cold dark matter as the material comprising such halos.

Jeremy Mould

Nuclear waste policy falls victim to early general election

London

THE British government's abandonment of plans to build a shallow low-level nuclear waste dump at one of four rural sites has thrown its nuclear waste disposal policy into disarray and prompted accusations of electioneering. The investigations, by the Nuclear Industry Radioactive Waste Executive (NIREX), began last October in the face of fierce local opposition; all four sites were in the constituencies of Conservative members of parliament, and at the time the decision was made the general election date of 11 June — announced on Monday this week — already looked a certainty.

The decision was announced by the Secretary of State for the Environment, Nicholas Ridley, on 1 May, the day after he received revised costings from NIREX, indicating that it would be almost as economical to dispose of low- and intermediate-level waste together in the same deep underground repository. NIREX denies that its recommendation to stop the tests was influenced by political considerations.

Ridley took the decision without consulting the government's own Radioactive Waste Management Advisory Committee, a group of independent experts who were closely involved with the choice of the sites and who had until last week been kept informed on the progress of the investigations. Members of the committee were angered by Ridley's failure even to give them prior notice of his decision, and

they regard his actions with scepticism.

Members of NIREX were themselves surprised at the speed with which Ridley approved and announced their recommendation, given that the investigations were due to be completed within days.

Initial estimates of the cost of shallow burial were around £125 per m³ but, according to NIREX, increased engineering safeguards, necessary to quell public anxiety, meant that the costs had risen to between £500 and £1,100 per m³; co-dumping with intermediate waste would provide a more publicly acceptable solution and would cost between £750 and £1,200 per m³ of low-level waste. The shallow repository would have come into operation in the early 1990s. Now, the 25,000 m³ of low-level waste produced each year will continue to be stored at the British Nuclear Fuels Ltd (BNFL) repository at Drigg, in Cumbria, and on site.

In recommending the abandonment of plans for a shallow dump, NIREX conceded that public acceptance of radioactive waste disposal is unlikely to be achieved in the short term.

Simon Hadlington

Australia abandons nuclear energy

Sydney

AUSTRALIA'S rejection of nuclear energy is complete with a change of name and direction for the Australian Atomic Energy Authority (AAEC). In its new guise as the Australian Nuclear Science and Technology Organization (ANSTO), it will no longer carry out research into uranium mining and enrichment or other aspects of the fuel cycle and power generation which were the basis of the AAEC's original charter. The only exception is SYNROC, its star project for nuclear waste disposal by immobilization in a ceramic.

The changes are intended to make the organization more responsive to the needs of industry. Besides basic research ANSTO will concentrate on the application of nuclear science to industry, medicine and the environment.

Australia came closest to adopting nuclear power in 1970 when plans were afoot to build a Canadian-designed power reactor 140 km south of Sydney. But plans were delayed as the anti-nuclear movement grew and large reserves of coal were discovered fairly near major cities. The project was shelved when a Labor government was elected in 1972, and no government has since revived it. Charles Morgan

West Germany strides towards CFC elimination by 2000

Munich

WEST Germany plans to push its chemical industry towards stricter controls on the production of chemicals that damage the ozone layer than those agreed at last week's meeting at Geneva of signatories of the Vienna convention. This emerges from the statement of an environmental official in Bonn on 4 May that the Geneva meeting did not go far enough towards eliminating emissions and production of chlorofluorocarbons (CFCs).

The West German Environment Office in West Berlin will soon begin talks with the chemical industry, asking for a near-total elimination of CFC production and emissions by 2000. The 31 states represented at Geneva in early May called for production cuts only of 50 per cent. This proposal will be put to an international conference in Montreal in September.

A government team of experts will begin a series of meetings with representatives of industry at the end of the month to discuss ways of reducing CFC production, starting with the aerosol industry. A member of the team, Hans-Jürgen Nantke, says that the aerosol industry will be asked to reduce CFC production to 50 per cent of current levels

within the next 18 months. If the industry refuses to comply, Nantke said, "we shall recommend that the Environment Minister issue regulations".

Surprisingly, the aerosol industry in West Germany is willing to abide by this request. Unlike CFC producers in Britain and France who strongly opposed the proposed reductions, German industry has already gone a long way towards elimination of all but 'essential' uses of CFCs. CFC production has been reduced from 53,000 tonnes in 1976 to 26,000 tonnes in 1986, with further reductions in sight. A spokesman for the aerosol industry attributed the fast pace of substitution of partially for fully halogenated compounds to pressure from scientists.

Fully halogenated CFCs such as CCl₃F and CCl₂F₂ are thought to contribute to the depletion of the stratospheric ozone layer. The effects of partially halogenated compounds such as CHClF₂ are believed to be less. Aerosol cans may also be replaced by mechanical pumps. But the non-flammable qualities of the fully halogenated compounds make them essential for applications in the medical technology and electronics industries.

Steven Dickman

Where to go now?

London

CRITICS have raised reservations about the ability of the Drigg site to cope with the extra volume of low-level nuclear waste beyond 1990 (see above). BNFL says that a new system of compaction will allow Drigg to be used into the next century.

For a long-term solution to the problem, NIREX is now concentrating its efforts on the search for potential deep sites. Investigations began in October 1985, and three options have been identified, an inland underground mineshaft, and beneath the sea bed with access from either an on-shore tunnel or an off-shore platform.

The British Geological Survey has identified several areas, including remote islands, that are hydrogeologically sound. Planning, transport and radiological studies are being made, and several potential sites should be announced within 12 months. Site investigations would then take up to three years and, allowing for a public inquiry and construction, the repository would not be in operation before 2003, according to NIREX estimates. □

Reorganization of Spanish science policy gets under way

Barcelona

AN important step has been taken towards setting in train the reforms in Spanish science policy under the Science Law approved last year (see *Nature* 324, 316-317; 1986). The latest decisions confirm the intention of the socialist government to build up a new and better orientated organization of science policy centred around the national plan for scientific research and development.

In recent weeks, the new structure for the organization of science policy has been published and appointments have been made to the principal posts. The Science Law assigns the preparation of the national plan for research to an inter-ministerial commission on which are represented the departments interested in

research. The key appointment is that of Dr Emilio Muhoz as secretary general of the plan, attached to the ministry of education and science.

Muhoz, a biochemist, has been director general of science policy since the socialist government came to power in 1982 and is one of those responsible for the new law. In addition to preparation of the plan, he will deal with other matters such as the coordination of international activities and will have several technical departments at his disposal intended to create an administrative structure to assist decision-making in science policy.

The new director general of science policy is Dr Luis Oro, an inorganic chem-

ist from the University of Zaragoza who was an official in the Comision Asesora de Investigacion Cientifica y Tecnica (CAICYT), the main granting body for research in the previous system.

An interesting new body created as part of the reorganization now under way is the national agency for evaluation and forecasting. This is meant to be an independent body for the evaluation of grant applications and is intended to continue the policy of the former CAICYT. The new agency will probably be based on the experience already acquired by CAICYT.

Pedro Puigdoménech

New UK technology centres launched

London

SEVEN regional technology centres are to be created in Britain, financed by a £1,548 million first-year budget provided jointly by government, industry and academic institutions to bring together the relevant partners interested in commercially exploiting technology.

A fund of £687,000 will be derived from government coffers while 22 universities, 14 polytechnics and 17 partners in industry, including GEC, Ferranti, Rolls-Royce, British Aerospace and British, will provide the rest. According to Mr Kenneth Baker, Secretary of State for Education and Science: "We are now pressing ahead with an initial seven centres which will build on existing successful schemes to transfer technological innovation from the laboratory into new products for business and industry. More centres will follow later... Whilst enabling business to capitalize on the opportunities created by new research developments, the close ties with education should foster a closer relationship between undergraduates, diploma and continuing education courses and the world of work."

The government funds, up to about £100,000 a centre for the first year, will help establish the network, with funding reducing in the second year as the centres move towards self-financing. The industrial and higher-education partners will provide the rest of the cash, equipment and personnel to operate the centres.

Typically, the centres will offer these services in manufacturing technology, biotechnology, computers and micro-electronics, information technology and materials.

Bill Johnstone

Germans build synchrotron

Munich

WEST Germany wants to fill a gap in the spectrum of facilities for "medium-energy physics" with a device called COSY. A cooled synchrotron, COSY is a proton accelerator that will be used to conduct experiments at kinetic energies of 2.5 GeV that higher-energy installations such as CERN (the European Organisation for Nuclear Research) have neglected. The results are expected to reveal more about phenomena such as the creation and decay of strange baryons and mesons.

COSY will be constructed at Kernforschungsanlage (Nuclear Research Institute or KFA) at Jülich and should be completed by 1992. Ninety per cent of the estimated cost of DM90 million will be assumed by the Federal Ministry for Research and Technology.

A storage synchrotron with phase-space cooling, COSY will offer a unique feature for such synchrotrons: a very narrow proton beam. Although it will not offer the same beam intensity as its French predecessor, the Saturn Ring at Saclay (10^{10} protons s^{-1} compared with Saturn's 10^{11} protons s^{-1}), COSY's beam width of about 0.5 mm will permit observations of various collisions. The Saturn Ring, by comparison, has a beam width of 3-4 mm.

Shaped like a racetrack, COSY will use one of the KFA's accelerators as an injector. Targets may be placed within the racetrack and outside it, and sensitive spectrometers are also available to observe the meson 'sparks' from each collision.

But a detector has yet to be built at the KFA that can measure the most energetic elastic collisions. Such a spectrometer, said scientific director Kurt Kilian, "will come sooner or later". Steven Dickman

International acid rain protocol hopes

London

AGREEMENT on the contents for a protocol on the so-called international acid rain treaty is coming closer, as representatives met this week in Geneva for another round of negotiations. A limit on emissions and deposition of nitrogen oxides is the goal of the United Nations/Economic Commission for Europe UN/ECE working party, set up as a result of the 1979 Convention on Long-Range Transboundary Air Pollution.

Previous meetings of the working party have been acrimonious, largely because of opposition to a nitrogen-oxides protocol by Britain and Eastern European countries. Now, however, the first proposals for a protocol have been agreed, centring on a phased approach to controls.

This week's negotiations will focus on measures to be taken during each phase, and the setting of "critical load targets" for nitrogen-oxide depositions. The concept of critical load — "the highest load which will not cause chemical changes leading to long-term harmful effects on the most sensitive ecological systems" — now seems generally accepted as a more scientific approach than the politically based 30-per-cent reduction set in a similar protocol on sulphur emissions. Critical load targets could be achieved by specifying a percentage reduction in emissions and transboundary fluxes, and/or by technical standards specifying measures to be taken.

Non-governmental organizations attending the meetings have been calling for an overall reduction of 75 per cent, with a limit on nitrogen oxides deposition of less than 10 kg per hectare per year.

Kathy Johnston

● Britain's Central Electricity Generating Board (CEGB) last week announced plans for a 10-year programme to install low nitrogen-oxide burners at 12 major coal-fired power stations. The retro-fitting programme, which is expected to begin in 2-3 years, will lead to a reduction in nitrogen-oxide emissions of nearly 30 per cent on 1980 levels, according to the CEGB. □

Doubts voiced in Japan over controversial AIDS legislation

Tokyo

JAPAN'S proposed AIDS (acquired immune deficiency syndrome) prevention law is coming under a barrage of criticism from doctors, lawyers and haemophiliacs.

The AIDS bill, due to be debated in the current session of the Diet, was hurriedly drafted by the Ministry of Health and Welfare earlier this year during a blaze of publicity surrounding the death of Japan's first female AIDS patient. But the proposed legislation was immediately condemned as "inhumane" by the National Association of Haemophiliacs (see *Nature* 326, 232; 1987), and the Japan Society of Blood Transfusion also opposed the law on the grounds that it would drive potential carriers underground. Now lawyer and doctors have joined the chorus of criticisms.

The most contentious provisions in the proposed legislation are that: doctors must report the age, sex and source of infection of AIDS-virus carriers to the prefectural government concerned within seven days and, if a doctor judges that a patient could ignore his instructions, he must report the patient to the prefectural authorities; prefectural governors are authorized to question and recommend or order AIDS tests and medical check-ups for people infected with AIDS or suspected of being infected; those who fail to keep the secrets of patients will be imprisoned or fined.

The Japan Civil Liberties Union, an association of lawyers and law scholars, on Saturday declared its opposition to the legislation because it may violate the human rights and freedom of patients and also because, in the union's view, public health specialists were not adequately consulted in the rush to draft the law — an opinion supported by a society of doctors

concerned with public health.

This week, the Japanese Society for Hygiene will send an appeal to the Ministry of Health and Welfare listing three major defects in the law: (1) the law was drafted very hurriedly without sufficient consultation, and without adequate explanation of intent; (2) the major purpose of the law is to identify the population of carriers, but the law may very well scare away at-risk groups, leaving haemophiliacs as the most visible potential carriers and exposing them to social discrimination; and (3) although the law includes punishment against violation of privacy, very few cases of violation of privacy of individuals have ever been prosecuted.

Evidence is already mounting to support the society's view. Doctors report a decrease of about 20 per cent in the number of haemophiliacs visiting AIDS clinics since news of the law broke. And, according to Yukio Yasuda, of the National Association of Haemophiliacs, a Japanese child infected with AIDS has been refused admission to pre-school, and other school required a pupil suffering from haemophilia to bring a certificate stating that he was not infected with AIDS.

Professor Gen Ohi of Teikyo University, who drafted the hygiene society's appeal, contrasts Japan's reaction to AIDS with that of the United Kingdom and the United States, where there are far more AIDS cases. Mr Kenneth Clarke, UK Minister of Health, rejected notification of AIDS in 1985, and at a meeting at the US Center's for Disease Control in February participants almost unanimously rejected mandatory testing. Instead of the present AIDS legislation, Ohi advocates setting up open-door clinics, where patients can receive treatment and tests in complete confidence. **David Swinbanks**

Reagan seeks expert advice on AIDS

Washington

PRESIDENT Ronald Reagan has announced a national commission to advise him on ways of dealing with the spread of AIDS. But commission members, yet to be announced, are not going to find it easy to reconcile the rapidly diverging views on public information about AIDS and AIDS testing.

The commission "will review research, assess the long-term impact on the health-care system, and recommend ways to protect those who do not have the disease". On the last topic, the administration appears deeply split between those, like Education Secretary William Bennett, who favour compulsory testing and stress on morality

in education, and those, like the Surgeon General C. Everett Koop, who favour expanding voluntary test programmes and public education on the use of condoms (see *Nature* 327, 3; 1987).

Backing for Koop's view comes this week in a new report from the Public Health Service's Centers for Disease Control (CDC). The report has not yet been made public but, according to CDC researchers, it argues that compulsory testing would be a use of resources that could not be justified by the current pattern of spread of the AIDS virus. Instead it recommends a major programme to encourage high-risk groups to come forward voluntarily for testing. **Alun Anderson**

Telescope opens

London

THE £20-million, 15-m-diameter joint venture James Clerk Maxwell sub-millimetre telescope, opened by the Duke of Edinburgh on the top of the 4,000-m Mauna Kea in Hawaii on 27 April, will start its observing career over-subscribed 6.5 times. The Royal Observatory, Edin-



burgh, which is looking after the programme, received 85 proposals of which it could only choose 15 for the first six-month observing period beginning in July. Optimum running time will be 16 hours a day, unless further money can be found for more extended operation. Theoretically, the telescope could work round the clock as the protective membrane (shown above — with its designer) effectively cuts out all solar radiation.

Angela Croome

WHO derides vaccine scare

London

REPORTS of a link between the successful global smallpox eradication programme and the emergence of AIDS have been dismissed by the World Health Organisation (WHO). Dr Jonathan Mann, director of the WHO AIDS programme, said that allegations that the smallpox vaccine, vaccinia, may have activated HIV infections "join many other unproven and speculative ideas about the origin of AIDS". Mann is anxious that unfounded speculation could jeopardize other vaccination programmes, and called for work that links smallpox vaccine with AIDS to be submitted for scientific scrutiny.

It was emphasized at the 40th WHO assembly meeting this week in Geneva that the world should concentrate on action to prevent the spread of AIDS, rather than speculating on its origins.

Kathy Johnston

● The British aid organization War on Want says up to 75 million Africans could soon be affected by AIDS. Using figures from the WHO conference, the group predicts that the virus will spread along Africa's lines of transport to Zimbabwe, Mozambique and Angola, and that South Africa "may face a major and widespread epidemic in the near future." □

El Niño behind penguin deaths?

SIR—In the South Atlantic Ocean in February 1986, large numbers of Rockhopper penguins (*Eudyptes crestatus*) were found dying along the coast of the Falkland Islands^{1,2}. By late May 1986, 3,000 penguins were found dead in one rookery. The two hypotheses put forward to explain this mortality of rockhopper penguins — competition for food with fishermen, and marine pollution from radioactivity^{1,3} — are unconvincing^{3,4}. Overfishing should not cause sudden, severe and widespread changes over a large area in the South Atlantic. The Gentoo penguin (*Pygoscelis papua*) and the magellanic penguin (*Spheniscus magellanicus*), which feed further inshore and often on fish were less affected, suggesting food was still available. And there is no evidence that radioactivity caused penguins to die.

In February 1986, while penguins were dying in the Falkland Islands, Cape cormorants (*Phalacrocorax capensis*) off the coast of Southwest Africa deserted their nests⁵ and magellanic penguins along the Argentine coast from Punta Tombo to Cabo Virgenes fledged chicks and moulted as usual. The small rockhopper colony at Isla Pinguino also appeared normal. It was not until April 1986, that excessive mortality and malnourished penguins began appearing. Thousands of penguins were found dead along the Argentinian coast of Santa Cruz. In April, only live birds were seen at Cabo Domingo but in May 276 rockhopper and 10 magellanic penguins were found dead. About 500 km north, at Cabo Blanco, 294 rockhopper penguins were found dead in the period 4–9 May. All were emaciated and most were moulting. Examination of eight of the most recently dead birds showed they had little or no fat, suggesting death by starvation. The contention that the penguins died from radioactivity leaking from sunken British submarines³ was not upheld by the background level of radioactivity found in their stomachs.

There is other evidence that suggests that the penguins died because food was unavailable. In Argentina, penguins were moulting in early May 1986, two to three months later than normal, as occurs when food is scarce or not available. Similarly, rockhopper penguins in the Falkland Islands, which normally moult and depart in April, were still present in late May. They may have been unable to gain sufficient weight to moult and, therefore, delayed both moulting and leaving. Furthermore there are several records of penguins being found north of their normal range in the South Atlantic in February to August 1986, suggesting birds wandered to other areas in search of food.

Presumably, the lack of food was

caused by a change in oceanic conditions. The locations and timing of penguin mortality shows a large area was affected for several months. The South Atlantic had stronger westerly winds (poleward from 40 degrees South) from December 1985 to September 1986, than the average 1980–1983 winds⁶. These strong winds should have intensified wind-driven inshore upwelling and benefited coastal species. Indeed, breeding magellanic and rockhopper penguins along the Argentine coast appeared normal. In the Falkland Islands, penguins fed closer to shore in 1986 than in 1985⁴. What could have caused abnormal oceanic conditions in the Atlantic in 1986?

In February and March 1986, the breeding success of seabirds in the Pacific was extremely low, as is typical during a El Niño Southern Oscillation (ENSO) in the Pacific Ocean^{7,8}. Massive die-offs and reproductive failure of seabirds are common during ENSO events in the Pacific. Before the ENSO events along the coast of Peru, the Southeast trades blew more strongly than normal and upwelling was stronger than usual⁹. Strong westerly winds in the Atlantic may be the critical factor in changing offshore conditions in the South Atlantic. In both oceans a strong wind-driven upwelling system is perturbed by unusually strong surface winds.

During 1986, atmospheric conditions in the South Atlantic were peculiar: the southern summer of 1986 was unseasonably long and hot and the Falkland Islands had an influx of warm water species¹. These changes are consistent with a large-scale warming. Mechanisms of atmospheric forcing in the tropical Atlantic analogous to the Pacific ENSO have been suggested and there are data showing the Benguela Current along Africa is affected¹⁰.

I should therefore like to suggest that the ultimate cause of the extensive mortality of penguins in the South Atlantic Ocean in 1986 was an ENSO event.

P. DEE BOERSMA

*Institute for Environmental Studies
and Department of Zoology,
University of Washington,
Seattle, Washington 98195,
USA*

1. Lyster, S. *Falkland Island Fdn Newsl.* 5, 2–4 (1986).
2. Rich, V. *Nature* 322, 4 (1986).
3. Rich, V. *Nature* 322, 296 (1986).
4. Bourne, W.R.P. *Nature* 322, 679 (1986).
5. Crawford, R.J.M., Williams, A.J. & Crawford, P.B. *S. Afr. J. Mar.* 4, 119–123 (1986).
6. *Climate Diagnostics Bulletins* (NOAA, Washington DC, January–September 1986).
7. Schreiber, R.W. & Schreiber, E.A. *Climate Diagnostics Bulletin* 7–8 (NOAA, Washington DC, March 1986).
8. Bowman, R.I. *Climate Diagnostics Bulletin* 8 (NOAA, Washington DC, April 1986).
9. Cane, M.A. *Science* 222, 1189–1195 (1983).
10. Shannon, L.V., Boyd, A.J., Brundrit, G.B. & Taunton-Clark, J.J. *mar. Res.* 44, 495–520 (1986).

Publishers needed

SIR—Wallace S. Broecker, a distinguished scientist in his own field, has like many others been led to the view that he can also solve problems in at least one other field — publishing (*Nature* 326, 207; 1987). Well ... I believe that Broecker's main argument is false. His error is not new. He says "I do not need a publisher because I can edit books cheaply by means of word processing, computer assistance for illustrations, etc...". Very true. Broecker can produce his book this way. The trouble is that publishers seek to design books that will appeal to readers as well as to the author, so that the readers buy it and so that it becomes their book.

This is why publishers desire to "dominate decisions about format, cover design, marketing and so on". It seems not to have occurred to the author that publishers may have some ideas about their trade. No doubt Broecker can also design a cover, provide the illustration and do the layout. I can also believe that he can fry eggs and buy bread, and sustain himself with this diet: he will probably produce a pamphlet next week denying the capacity of cooks and the ability of nutritionists in *Home-cooking for the Gourmet*. Similarly with art: when you have seen one portrait, you have seen them all.

Publishing like any other industry expects a profit. I do not understand why Broecker thinks it wants a profit "above all". His "financial outlook" assumes that all the books will be sold: with this assumption, any financial outlook looks good.

But on one point I object very strongly: publishers should not need the help of "grants from charitable foundations" or (I might add) government subsidies, direct or indirect, to publish books. If Broecker expects "above all" to make a profit of \$142,000 with one book, he has to take some risk.

PHILIPPE BOULANGER

*Pour la Science,
8 rue Férou,
75006 Paris, France*

But not female

SIR—Your news article on the King Faisal International Prizes (*Nature* 326, 118 1987) contains the statement "On the other hand, of 14 science or medicine prize-winners so far, none has been female or Jewish". That is not correct. I am a co-winner of the King Faisal International Prize in Medicine (1984) and I am Jewish, whether or not it was known to the officials.

MICHAEL FIELD

*Department of Medicine,
College of Physicians & Surgeons
of Columbia University,
New York, New York 10032, USA*

Quantum information storage

Whether energy is necessarily dissipated in computer operations remains an open question. The development of optical computers should help decide the question.

ONE of the rich but unsettled questions of computer science is how to tell what is the least amount of energy required for the storage of a single binary bit of information in a memory device. There are even some who hold that there may be no lower limit different from zero. For while it is plain that work must be done and energy expended to change the memory state of a macroscopic storage device, say a ferrite core, the amount of energy required is very much a function of the size of the storage devices, which are steadily shrinking in dimensions.

If dreams come true and the dimensions of storage devices are limited only by considerations of quantum mechanics, might not the irreducible energy cost also shrink to zero, at least if it is recognized that energy can be recovered when ordered storage elements are allowed to revert to a random condition?

That is almost certainly an over-facile argument, as is suggested by what is known of the thermodynamics of information storage. Stored information not merely has, but is, negative entropy. Shannon showed that half a century ago, but the analogy between the operation of storage devices and classical thermodynamics can be made much more direct.

The cycle of storing a bit of information and then retrieving it is very like the cycle of operations in an ideal heat engine—the Carnot cycle. The bit is stored, by the expenditure of energy, and then retrieved, whereupon some energy may be recovered. It is readily appreciated that the efficiency of the process will be a function of the temperature, if only because the chance that the stored bit of information will be corrupted, or lost, by random processes in the storage element which are strictly negligible only at absolute zero.

So why not calculate just what amounts of energy are involved for some real device? Hitherto, the impediment has been the analogy in the field of information of the departures from ideality of real-world heat engines. Ferrite cores, being essentially magnets whose direction of magnetization can be changed, are usually outrageously much larger than would be required for the storage of a single bit. The work done in taking them through a cycle may be accurately measured by the area of the usual hysteresis loop, but the quantity thus obtained bears on the size of the ferrite core, not on the energy entailed in storing and retrieving a

single bit of information.

That is why there will be much interest in a microscopic information storage device designed (as yet only theoretically) by a group at the University of Stuttgart, G. Obermayer, G. Mahler and H. Haken (*Phys. Rev. Lett.* **58**, 1792; 1987). They start from what is known of the construction of microelectronic devices by epitaxial growth in which layers of difference semiconductors alloys a few Ångströms thick may be accumulated in a multi-decker sandwich. If the electronic properties of the successive layers are suitably chosen, some of them may serve as traps for electrons, or for electron holes, which are the best realizations yet of the classical problem of elementary quantum mechanics, that of a particle in a box (whence the name 'quantum well' structures).

Obermayer *et al.* have worked (on paper) with a three-layer sandwich, one of which is pure gallium arsenide (GaAs) and the other two are alloys of that with different proportions of aluminium. By a suitable choice of the proportions of aluminium, it should be possible to make such a device so that it functions as two separate quantum wells separated physically by a few nanometres. (The calculations are based on a device in which the wells are each 4-nm thick and separated by 16 nm.)

The accessible electronic states in each of the two wells are listed easily enough. First, there are the tightly bound electron states which together are called the valence band, which are ordinarily full, at least in a semiconductor with a finite band-gap. An electron missing from such a state appears as a 'hole', that is as a positively charged entity whose apparent mass is much greater than that of an electron. Then there are the relatively free electron states collectively constituting the conduction band, into which valence electrons may be promoted with increasing temperature, but which otherwise acquire occupants by conduction from elsewhere.

The calculations of the energy states of this system is complicated, and has to be numerical, but it seems clear that each of the wells acts as a trap for electron holes, which are totally confined therein, if for no other reason, because of their great mass relative to that of an electron. The device is therefore one that can store information, as represented by the heavy-electron holes in each well, at least so long

as a hole is not filled by the arrival of a stray electron from elsewhere.

In principle, the higher-energy conduction band is therefore the chief means by which information can be communicated from one well to the other, and is thus the means by which holes can be switched from one well to the other. This, of course, has to be accomplished by electron excitation, most simply by light absorption. Obermayer and his colleagues show that it is indeed possible to prepare particular states of the system in a systematic way; suitable potential biasing of the two wells will allow each of them to be driven by light (usually in the infrared) of different frequency. One of the remarkable properties of the general conclusion is that states of the system once prepared are stable at least on the timescale of thermal degradation, while switching speeds may be measured in nanoseconds, all of which is good news for those who hope to use GaAs quantum wells as the elements in optical computers.

The philosophical interest of the system is that, in principle, it makes it possible to consider more completely than has previously been possible the thermodynamics of an information storage device. Obermayer and his colleagues note that many of the muddles of the past have arisen because people have confined themselves to isolated systems, 'closed' in the sense of thermodynamics. Their device, which they figuratively embed in a larger block of GaAlAs, is open (in the same sense) both to electrical influences and because of the interaction of the electron states both with the surrounding radiation bath and with the vibrations of the lattice structure. As a consequence, it should be possible to examine in the detail required the extent of which energy is dissipated in such a device not merely by random interactions with the outside world, but also by the imprecisions of the system for preparing particular storage states.

In an ideal world, it should be possible to set about the measurement of the quantities concerned. The obvious difficulty, made more apparent by these calculations, is that the properties of quantum-well systems are certain to be sensitive to the fine details of their construction. Dimensions will have to be exact, as must be the chemical composition of the several layers of the GaAs sandwich. It remains to be seen how successful these enterprises will be.

John Maddox

Genetic probes

DNA fingerprints applied to animal and bird populations

William G. Hill

THE DNA probes developed by Alec Jeffreys and colleagues¹, which hybridize to hypervariable minisatellite regions of human DNA to produce DNA fingerprints, are becoming important in forensic and parentage studies in man. These probes, constructed from single tandem repeats of a short human core sequence, have recently been shown to reveal similar hypervariability in mice², dogs and cats³. On pages 147 and 149 of this issue, Wetton and colleagues⁴ and Burke and Bruford⁵ show that DNA fingerprints of birds can be obtained in the same way, and use them to demonstrate mixed parentage among nestlings of house sparrows. The minisatellite probes provide a powerful technique that is likely to be widely used in family, demographic and linkage analyses of many wild and domestic species.

In studies of the breeding structure of wild populations it is necessary to establish relationships, but with the markers previously available in house sparrows, for example, only about one-half of cases of non-paternity could be established⁴. The identification of parentage is also important in domestic species where large sums of money are paid for inseminations or where animals are being progeny tested. In cattle, for example, this need has been met by highly polymorphic blood-group markers, but for most species antisera are not available or loci are not sufficiently polymorphic. It now turns out that the DNA probes (see figure) obtained in man by Jeffreys and colleagues are effective markers in most other species examined.

Using only one of the probes (33.6), Wetton *et al.* identify about 60 bands, each representing a DNA segment of a different length that hybridizes to the core sequence, in the house sparrow. Although the numbers of birds examined are small, estimates of the same order of magnitude have been obtained in both studies reported in this issue^{4,5} for the probability that birds will have the same pattern. These probabilities are very low (see table), and are still so for full sibs, enabling close relatives to be distinguished. In a family of sparrows of two adults and eleven offspring raised in four broods in a closely monitored population, Burke and Bruford⁵ find one bird that has six bands absent from either parent. This result clearly indicates that the bird is not an offspring of the mating, and is almost certainly a maternal half-sib of the rest. Similarly, Wetton *et al.*⁴ find that a male

from an adjacent nest box is probably the father of one of an otherwise full-sib family.

The minisatellites are not the only new method being used for such population studies: recently, Quinn *et al.* reported⁶ the use of DNA restriction-fragment length polymorphisms (RFLPs) in paternity/maternity analyses of lesser snow geese in Canada. These authors already had evidence for either extra-pair mating or egg dumping because blue offspring

33.6 (A G G G C T G G A G G)₃₁₈

33.15 (A G A G G T G G G C A G G T G)₂₉

The minisatellite probes consist of tandem repeats of closely related variants of these consensus sequences, the repeat unit in 33.6 being a diverged trimer¹⁰.

had been found from pairs of white birds, white being recessive. Their analysis confirms this idea, and reveals other extensive illegitimacy (6 out of 17 birds of a selected group). Because the heterozygosity of the probes is low, the power of these RFLPs, which presumably reflect base differences or insertions, to detect non-paternity seems substantially below that of the minisatellites. Also, Quinn *et al.* had to construct new probes and hybridize each one, whereas the two minisatellite probes of Jeffreys and colleagues seem relatively ubiquitous.

Jeffreys and Morton find that among eight different breeds of dogs and five of short-hair domestic cats, the level of band sharing is higher than in sparrows (see table) or in man, but there is nevertheless substantial variability³. The only animals of the same breed analysed, two whippets, are not significantly more similar to each other than to dogs of different breeds, but it would be surprising if a more thorough study revealed no breed structure.

Analysis of laboratory animals can provide substantial information about the minisatellites. Jeffreys and Morton also investigated³ the inbred lines C57BL/6J (B6) and DBA/2J (D2) and the B × D

recombinant inbred strains, established from the F₂ hybrid of the two-way cross about 60 generations previously. These results give the first evidence of X-linked inheritance, as an invariant region shows a lower hybridization signal in males. Most fragments in B × D can be traced to one or other progenitor, and those that are absent from B or D are present in only one recombinant strain. This result illustrates mutation at these regions, observed also in man¹ and in birds⁵, but at a lower rate ($<0.5 \times 10^{-4}$ per kilobase minisatellite per gamete) than in man (10^{-4}). One site, however, is highly variable in length over the lines, corresponding to a mutation rate in excess of 15×10^{-4} per kilobase minisatellite.

With DNA fingerprints, in contrast to blood groups, electrophoretic loci or conventional RFLPs, allelism cannot be established simply because so many bands can be seen on a single gel. Without allelism, although a mating pair can be excluded as parents of an individual, specific exclusion of the putative mother or father is not positive but relies on probability calculations. Demographic analyses of widely divergent populations may not be practicable because fragments of the same mobility are not necessarily isoallelic: linkage data for markers with, say disease loci from different populations, and indeed families, cannot then be combined.

The recent studies of mice² and of birds⁵ show clear indications of allelism and, where established, alleles segregate in the usual mendelian ratios. The association with known markers enables the chromosome location to be assigned to eight of the thirteen markers identified in mice². Some loci produce multiple bands or are so closely linked that they cosegregate^{2,5}, so care must be taken in calculating probabilities of identity.

Although probes such as 33.6 and 33.15 locate many sites, they can be used to find new probes that hybridize only to single loci. In one such study on man, in a sample of 79 individuals only 3 were homozygous and 77 alleles were identified⁷. This approach has been developed by R. White's group, which reports on the use of 77 different probes that identify variable numbers of tandem repeats⁸. The average heterozygosity of 67 of these repeats with 3 or more alleles was more than 70 per cent. Providing the loci are dispersed over the genome, they will augment, and

Genetic analysis using minisatellite probes 33.6 and 33.15

	Probe	House-sparrow	Dog	Cat
Mean number of bands per individual	33.6	6	16	8
	33.15	15	19	13
a		0.22	0.46	0.47
b		6×10^{-34}	2×10^{-21}	3×10^{-14}
c		4×10^{-18}	5×10^{-10}	7×10^{-7}

Probability: (a) that unrelated individuals have the same band; and that two unrelated individuals (b) and two full sibs (c) have the same band pattern for both probes (from refs 3,5).

probably replace, conventional RFLPs with low heterozygosity in linkage studies.

Presumably, locus-specific probes can also be developed for animals and birds, and may be of value in establishing linkages to genes affecting important traits such as growth rate and milk production. This may be useful in breeding programmes⁹, and perhaps may help to locate such genes and to elucidate their mode of action. □

1. Jeffreys, A.J., Wilson, V. & Thein, S.L. *Nature* **314**, 67-73 (1985).

2. Jeffreys, A.J., Wilson, V., Kelly, R., Taylor, B.A. & Bulfield, G. *Nucleic Acids Res.* **15**, 2823-2836 (1987).
3. Jeffreys, A.J. & Morton, D.B. *Anim. Genet.* **18**, 1-16 (1987).
4. Wetton, J.H., Carter, R.E., Parkin, D.T. & Walters, D. *Nature* **327**, 147-149 (1987).
5. Burke, T. & Bruford, M.W. *Nature* **327**, 149-152 (1987).
6. Quinn, T.W., Quinn, J.S., Cooke, F. & White, B.N. *Nature* **326**, 392-394 (1987).
7. Wong, Z., Wilson, V., Jeffreys, A.J. & Thein, S.L. *Nucleic Acids Res.* **14**, 4605-4616 (1986).
8. Nakamura, Y. *et al. Science* **235**, 1616-1622 (1987).
9. Soller, M. & Beckmann, J.S. *Proc. 2nd World Congr. Genet. Appl. Livestock Prod.* **6**, 395-404 (1982).
10. Jeffreys, A.J. *et al. Nature* **316**, 76-79 (1985).

William G. Hill is Professor of Animal Genetics at the University of Edinburgh, West Mains Road, Edinburgh EH9 3JN, UK.

Earthquake prediction

Studies of San Andreas fault

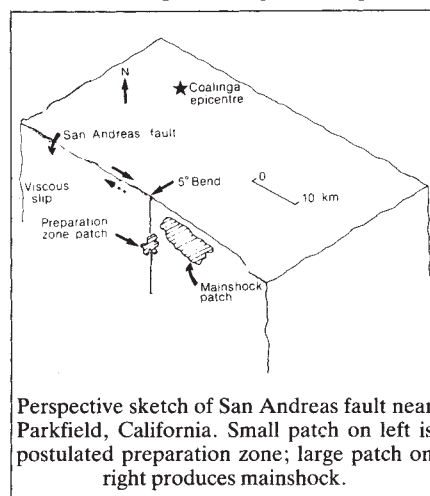
William D. Stuart

ONLY ten years ago, geophysicists studying earthquake zones were content if they could link observed deformation of the ground surface with slip at depth on a fault plane. Basic problems, like the precise conditions that trigger seismic faulting and the mechanical causes of spatial and temporal variation in seismicity, were left unexplained. These problems, previously only loosely connected, are now closer to resolution because of new observations and theoretical results. The main reasons for the reconciliation of data and mechanical theory are the recognition that specific areas of crustal fault planes are either dominantly brittle or dominantly ductile, and that different areas of one or more faults interact continuously.

One clear example of fault interaction, involving the San Andreas fault near Parkfield, California, is described by Poley *et al.*¹ on page 134 of this issue. They report simultaneous decreases of seismicity (magnitude > 1.6) and of surface fault creep (slip) starting just after the magnitude 6.7 Coalinga earthquake in May 1983. The seismicity and the nearby creep rate resumed after 15 months, whereas the creep rate farther away recovered after 27 months. The Coalinga epicentre was 30 km north-east of the north-west-trending San Andreas fault and was nearly opposite the section of anomalous creep and seismicity.

A fault model presented by Simpson and co-workers² at a recent conference on intermediate-term earthquake prediction accounts for the main features of the creep and seismicity anomalies. The model divides the area of the San Andreas fault into numerous interacting dislocation surfaces, each having a specified fault rheology. In the model, slip of the Coalinga earthquake creates shear stresses that subtract from ambient shear stresses on most of the San Andreas fault near Parkfield, thus accounting for the observed decrease or reversal of fault creep rate and the silencing of seismicity. The Coalinga

earthquake also alters compressive stresses, but the consequences are less apparent. The contribution of Simpson *et al.* is not so much a confirmation that fault interaction exists³, but rather the demonstration that where the San Andreas fault is not brittle, it is to a good approximation sliding in a viscous fashion. In essence, slip of the Coalinga earthquake imposes a



stress pulse on the San Andreas fault, and the ensuing relaxation is manifested by delayed recovery of creep and seismicity to normal rates.

The Parkfield section of the San Andreas fault is also of special interest, of course, because a moderate earthquake (magnitude 5-6) is expected there in the next few years based on the observed 21 ± 3 year recurrence of such events⁴. The last moderate Parkfield earthquake was in 1966 and its epicentre was near a 5° bend of the fault trace (see figure), with subsequent seismic rupture extending 20 km to the south-east. Poley *et al.*¹ suggest that the 'Parkfield Preparation Zone', a fault area 8 km long and 12 km deep, including the 1966 focus and the zone of anomalous creep and seismicity discussed above, is unusually sensitive to stress perturbations generated elsewhere: it is a kind of

tectonic amplifier. As evidence, Poley *et al.* note that observed fault creep above the preparation zone is even more anomalous than predicted by the theory of Simpson *et al.* The hope is that signals precursory to the next moderate earthquake will be similarly amplified.

Poley *et al.* do not give a precise mechanical description of the preparation zone, but they suggest that it may have distinctive geometric or rheological properties, related perhaps to the fault bend. In fact, it seems likely that the preparation zone is just a small brittle area or patch of the fault as in the figure. This patch, surrounded by slipping parts of the fault that concentrate stress there, but which resists slippage itself, would have its failure delayed whenever shear stress decreased, as by the Coalinga earthquake. Parts of the irregular patch could fail seismically from time to time, producing the observed background seismicity. Eventually, the entire patch might fail seismically in league with the larger brittle patch to the right in the figure to make a moderate earthquake like the 1966 event.

The loading and unstable (seismic) failure of such brittle patches have been theoretically studied⁵. In general it is found that crustal rigidity, patch size and patch weakening rate with fault slip determine the conditions for unstable failure. All theoretical models also indicate that the average rate of fault slip increases before instability, suggesting a way to anticipate future earthquakes at Parkfield. Few field data are available to check this result.

From the above model studies each of which focuses on a subset of fault properties and geometries, one can compose a general conceptual model for a tectonically active region. The main parts of the generic model are steady shear forcing from afar, fault geometries, and division of fault planes into brittle and ductile areas. The remote shear forcing causes ductile areas of faults between patches to slip more or less steadily, loading the temporarily locked brittle patches. Brittle patches eventually fail seismically and, when they do, their seismic slippage and stress drop act as an impulsive loading to all parts of all faults. Thus one can view a tectonically active region as an irregularly loaded assembly of jostling elastic blocks bounded by a network of interacting faults. It appears that it is possible to construct such models for different tectonically active parts of the world where sufficient field data are available to constrain fault geometry and rheology, and to solve numerically the associated boundary-value problems.

This class of models, however, goes only halfway because they do not explain the location of faults or why some areas of faults at comparable depths are brittle, yet others ductile. I suggest that the under-

lying explanation of these phenomena may be the strain field set up by different rock bodies because they have unequal rigidity and are immersed in a regional strain field. This idea is just the geological counterpart of the materials science problem of inclusions in a matrix. If the inclusion conjecture is correct then one can look forward to predicting faulting and seismicity using a mechanical model derived from a geological map. □

Plant ecology

Distribution of mycorrhiza throughout the British flora

Peter D. Moore

NO STUDY of plant ecology, particularly one involving factors affecting distribution, the availability of nutrients from the soil or the competition between individuals and species for limited inorganic resources, can afford to ignore the importance of mycorrhiza. These intimate associations between fungi and plant roots in the soil form a vital part of the physiology and life history of many plant species. They can even provide channels of movement of materials between plant individuals and between species, thus complicating studies of resource allocation and population ecology in the field. That they are widespread in nature has been appreciated for some time, but their abundance and ecological importance has been fully recognized only during the past 20 years. Information about the occurrence and type of mycorrhiza in different plant species is diffuse throughout the literature, but Harley and Harley have now collated¹ an exhaustive list of references to these associations in the British flora. This list both provides an opportunity to analyse the relative abundance of mycorrhiza and reveals some taxonomic and ecological peculiarities.

The list covers only the flora of the British Isles, but includes pteridophytes, gymnosperms and angiosperms. Harley and Harley have examined a significant proportion of the flora within each of these groups for the presence of mycorrhiza. In the pteridophytes, for example, they examined 82 % of the 74 British species, 72 % of which have mycorrhiza. Of the 13 gymnosperms native or well-naturalized in Britain, all have mycorrhiza. Among the angiosperms, the proportion examined is somewhat lower, but the general preponderance of fungal associations is still evident. In the dicots 53 % of the flora of 1,404 species were surveyed and 80 % of them have mycorrhiza, whereas in monocots 56 % of the 508 species are 76 % mycorrhizal. Thus, 54 % of British angiosperm species have

1. Poley, C.M., Lindh, A.G., Bakun, W.H. & Schulz, S.S. *Nature* **327**, 134-137 (1987).
2. Simpson, R.W., Schulz, S.S., Dietz, L.D. & Burford, R.O. in *Proc. Conf. on Physical and Observational Basis for Intermediate-Term Earthquake Prediction* (US Geological Survey, Open-File Report, in the press).
3. Mavko, G.M. *J. geophys. Res.* **87**, 7807-7816 (1982).
4. Shearer, C. U. S. *Geological Survey Open-File Report*, 85-754 (1985).
5. Rice, J.R. *Pure appl. Geophys.* **121**, 443-475 (1983).

William D. Stuart is at the United States Geological Survey, 525 South Wilson Avenue, Pasadena 91106, USA.

been inspected and 79 % of those are infected with mycorrhiza. The survey reveals that it is the non-mycorrhizal elements in the British flora that are exceptional, rather than the other way round.

Occurrence of mycorrhiza in some major flowering plant families in the British Isles (data from ref.1).

	Total species	Species examined (%)	Mycorrhizal species (% of those examined)
Dicots			
Ranunculaceae	45	56	96
Cruciferae	98	40	46
Caryophyllaceae	83	49	30
Chenopodiaceae	34	44	53
Geraniaceae	20	50	90
Leguminosae	87	47	90
Rosaceae	87	54	96
Onagraceae	21	71	100
Umbelliferae	70	47	79
Euphorbiaceae	18	56	100
Polygonaceae	41	51	38
Salicaceae	34	79	100
Ericaceae	25	84	100
Primulaceae	20	75	100
Gentianaceae	16	50	100
Boraginaceae	34	47	75
Labiatae	63	49	90
Plantaginaceae	8	88	100
Campanulaceae	14	64	78
Rubiaceae	24	50	92
Compositae	160	61	96
Monocots			
Liliaceae	41	71	100
Juncaceae	39	49	63
Iridaceae	15	33	100
Orchidaceae	53	92	100
Cyperaceae	107	54	31
Gramineae	173	60	87

Of the various types of mycorrhiza that have been described, it is the vesicular-arbuscular forms that are most frequent within the British flora. These forms are caused by a relatively small number of genera within the fungal family Endogonaceae, but are represented by many species. These aseptate fungi may form a hyphal coil within a cell, or produce a brush-like, branched haustorium (an arbuscule) as well as fat-containing, multinucleate vesicles with thick walls. The

vesicular-arbuscular form of mycorrhiza receives its name from these two types of structure.

Read^{2,3} has emphasized the relationship between mycorrhizal type and habitat. In low-latitude and low-altitude grassland and deciduous forest, where nitrogen is usually available in the form of nitrates, the vesicular-arbuscular endomycorrhiza predominates. At high latitudes and altitudes, in acid peaty habitats and in coniferous forests, nitrogen may be available only in a reduced, ammonium-ion form and in these environments, ectomycorrhizas and ericoid endomycorrhizas are the most frequent forms. These tendencies are apparent in Harley and Harley's taxonomic lists when examined from an ecological point of view. Families such as Ericaceae, Monotropaceae and Empetraceae almost all contain distinctive ericoid endomycorrhizas and ectendomycorrhizas. The lack of mycorrhizas in aquatic plants is also evident. Nymphaeaceae and Menyanthaceae have no mycorrhizal representatives, and Cyperaceae has only 31 % mycorrhizal species among the members investigated.

The close dependence of some families on mycorrhiza has long been known⁴, for example, the Orchidaceae and Gentianaceae, but other families are also surprisingly uniform in their fungal associations. Ranunculaceae, Geraniaceae, Leguminosae, Onagraceae, Labiatae and Rosaceae are all more than 90 % mycorrhizal in the species examined (see table). Yet other families have a surprisingly low proportion of mycorrhizal species: Caryophyllaceae, Crassulaceae and Polygonaceae have fewer than 40 %.

There is still much to be learned about the ecological advantages or disadvantages of being mycorrhizal. Such knowledge has already been of considerable economic value for forestry management in temperate and boreal regions. It is now becoming an important area of research in the tropical savanna grasslands, where increasing numbers of grasses are turning out to be mycorrhizal⁵. The productivity of grasses in these important pastoral areas could well be linked to the rapid development of an efficient mycorrhizal system early in the wet season. The current interest in the ecology of temperate mycorrhizal associations could provide an important springboard for the development of mycorrhizal research in tropical agronomy. □

1. Harley, J.L. & Harley, E.L. *New Phytol.* (Suppl.) **105**, 1-102 (1987).
2. Read, D.J. in *The Ecology and Physiology of the Fungal Mycelium* (eds Jennings, D.H. & Rayner, A.M.) 115-240 (Cambridge University Press, 1984).
3. Read, D.J. *Can. J. Bot.* **61**, 985-1004 (1983).
4. Newman, E.I., Child, R.D. & Patrick, C.M. *J. Ecol.* **74**, 1179-1183 (1986).

Peter D. Moore is Reader in Ecology at King's College London, (KQC), Campden Hill Road, London W8 7AH, UK.

Cancer biology

Oncogenes and cell growth

Peter Newmark

IF THE arrival of two journals* devoted to the subject marks the coming of age of oncogene research, a recent meeting[†] illustrated how age has a way of clouding the simple notions of youth. But, as so often in biology, new notions can flow out of unsolved problems.

Perhaps the most interesting new theme of the meeting was that normal cells, or their products, can exert control over those of their number that have been transformed into cancer cells. The control mechanisms are part of a complicated network of positive and negative controls on cell proliferation that is in continual operation. Presumably, the network is usually able to hold in check any cell in which an oncogene has arisen. But occasionally circumstances arise in which a cell escapes the net to proliferate in the uncontrolled fashion that produces a tumour.

It will take much effort to unravel the network, and particularly the inhibitors of cell proliferation, but several intriguing results and systems were described at the meeting. For example, if the viral *myc* oncogene is introduced into quail myoblast cells, they continue to proliferate rather than undergo differentiation to myotubes, the next stage in muscle formation. But differentiation is restored by the co-cultivation of the myoblasts with mammalian fibroblasts or normal quail myoblasts (F. Tato, University of Rome). In a poster, the same group extended the phenomenon to the suppression of the formation of 'foci' of *myc*-containing mouse fibroblasts (NIH 3T3 cells) when they were co-cultivated with mouse fibroblasts of a different line (C3H10T1/2).

A similar phenomenon was reported for keratinocytes provided with the viral *ras* oncogene (G.P. Dotto, Yale University). Grafted onto mouse skin, these cells result in the rapid development of carcinomas. But if the keratinocytes are mixed with fibroblasts from normal skin dermis the development of carcinomas is inhibited. Again, if the keratinocytes are injected into nude mice (whose immune system is impaired) tumours develop, whereas no tumours appear if the injected keratinocytes are first mixed with dermal fibroblasts. Thus, *ras* activation in a skin cell may not on its own be sufficient to initiate a carcinoma largely because of the inhibitory influences of adjacent cells.

By what mechanism are these 'influences' mediated? Tato reported that it is

necessary for the normal cells to be continually present in his system for their inhibitory effects to be felt; the medium in which they had been growing cannot mimic them. Consequently he argues that contact between the two cell types is necessary for inhibition. But secreted peptides can also inhibit the proliferation of some cells, and a failure of such cells to respond to the inhibitory peptides, which include interferons, may be a step in the direction of tumorigenesis (A. Kimchi, Weizmann Institute, Rehovot). In her latest experiments, Kimchi finds that the addition of an interferon antibody to the

medium in which rat kidney cells containing the *ras* oncogene are growing results in transformation of the cells, much as if a cooperating oncogene had been placed inside the cell. This experiment further provokes the notion that autocrine inhibition is an important mechanism in the normal control of cell growth and that disruption of the mechanism, by loss of either sensitivity to the secreted peptide or the ability to secrete it, can be a factor in tumour development.

Another factor can be the action of tumour promoters, chemicals that can push an initiated cell along the path to tumorigenicity. Phorbol esters are the favourite for study and the new focus of attention is on the gene transcription factors that mediate the induction of certain genes by TPA (12-*O*-tetradecanoyl phorbol-13-acetate) (B. Wasylyk, CNRS Strasbourg; P. Herrlich, Kernforschungs-

240-K superconductivity affirmed

FOR the past two months, rumours have been circulating about a phase in the Y-Ba-Cu-O system that superconducts at temperatures of ≥ 200 K. Several groups have seen drops in resistance at such temperatures, but the behaviour has not been reproducible, and many doubted that true superconductivity was being observed — a view reflected by the fact that the observations were reported mainly in newspaper articles rather than preprints. This week, superconductivity at ≥ 200 K finally attains scientific respectability, with the publication in *Physical Review Letters* of a paper by J.T. Chen *et al.*¹, reporting both resistance drops and the reverse a.c. Josephson effect in mixed-phase Y-Ba-Cu-O samples at temperatures near 240 K.

The phase responsible for superconductivity at ~ 90 K in the yttrium-based ceramics has the composition $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ (see, for example, ref. 2). Samples far from this composition contain more than one phase, and electron microscopy reveals a granular texture with a grain size of a few micrometres. It is only in these mixed-phase samples that resistance drops near 200 K have been observed: Chen *et al.* report a decrease of ~ 40 per cent between 240 and 200 K, followed by a flattening and then a sudden drop starting at 90 K to zero resistance at 60 K. As Chen *et al.* say, resistance measurements alone cannot confirm that the drop at 240–200 K is caused by superconductivity, because unless there is enough of the proposed 240-K superconducting phase to form a connected path, the sample cannot exhibit zero resistance.

Chen *et al.* took advantage of the granular nature of the sample, by looking for the a.c. Josephson effect. If there is a 240-K phase, then at temperatures between 240 and 90 K the sample may be thought of as comprising numerous coupled Josephson junctions — sandwiches of insu-

lating oxide (90-K phase) between two layers of superconductor (240-K phase). If the insulating layer is thin enough, a phase difference between the electron-pair wavefunctions on the two sides of the junction causes electron pairs to tunnel through the insulator, giving rise to a supercurrent. When there is a voltage across the junction, the phase difference increases with time, so that the current oscillates back and forth across the insulator; this is the a.c. Josephson effect. What Chen *et al.* observed was the reverse a.c. Josephson effect; that is, they used an alternating current to induce a constant voltage. They argue that this voltage can be distinguished from the time-averaged finite voltage that would result from rectification by diodes at grain boundaries by the observation (easily made on an oscilloscope) of a component that is constant with time. The induced d.c. voltage is observed starting at 240 K, coincident with the start of the resistance drop.

Why are the resistance measurements so irreproducible? Chen thinks the repeated cycling between room and liquid-nitrogen temperatures causes differential expansion and contraction of the grains, disrupting the electrical connections between them. Or, with time, oxygen may diffuse along grain boundaries, destroying the stoichiometry of the 240-K superconductor. The next step will be to identify and isolate the 240-K phase — only then will we know whether it is inherently unstable. Some purification will also be necessary if the four criteria of Tanaka — elucidation of the structure, observation of the Meissner effect, observation of zero resistance and experimental reproducibility³ — are to be satisfied.

Laura Garwin

* *Oncogene* (Macmillan, London); *Oncogene Research* (Harwood Academic, New York).

† *Oncogenes and Growth Control* European Molecular Biology Laboratory, Heidelberg, 26–30 April 1987.

1. Chen, J.T., Wenger, L.E., McEwan, C.J. & Logothetis, E.M. *Phys. Rev. Lett.* (in the press).

2. Rao, C.N.R., Ganguly, P., Raychaudhuri, A.K., Mohan Ram, R.A. & Sreedhar, K. *Nature* 326, 856–857 (1987).

3. Miura, N. *Nature* 326, 638–639 (1987).

zentrum, Karlsruhe; J. Piette, Institut Pasteur, Paris). Among the gene enhancer sequences to which these transcription factors bind are those of polyoma virus, a collagenase and the *c-fos* gene, and a good deal of effort will now be devoted to the question of whether and how TPA induces the factors and the relevance of these observations to tumour promotion.

The enhancers and transcription factors of the *c-fos* gene are themselves a subject of intense study, not least because the gene is one of the first to be turned on (transiently) when quiescent cells are stimulated by growth factors. The transcription factor that binds to a 20-base-pair palindromic enhancer sequence of the gene has been purified to homogeneity (R. Treisman, MRC Laboratory of Molecular Biology, Cambridge). Other *fos* research concentrates on the product of the *c-fos* gene and of its tumour-producing viral counterpart, *v-fos*. The central sequence of the *v-fos* protein seems to be the most important for its function and amino acid 138 has been identified as the key to its cell immortalizing effect. The lack in *v-fos* of the carboxy-terminal sequence of *c-fos*, within which is contained a negative regulatory region, is a key to the ability of *v-fos* to transform cells (R. Müller, European Molecular Biology Laboratory, Heidelberg). Deregulated *c-fos* can also transform cells; experiments with transgenic mice show that it eventually results in the appearance of bone tumours and not just the bone deformities reported at earlier times (E. Wagner, European Molecular Biology Laboratory).

The *fos* gene product is one of a group that seems to act in the nucleus, presumably on subsets of genes. Another is the product of *erb-A*. The proto-oncogene *c-erb-A* encodes the thyroid hormone receptor but the product of *v-erb-A*, the viral oncogene derived from *c-erb-A*, has lost the ability to bind thyroid hormones because of both a carboxy-terminal deletion and point mutations (B. Vennström, European Molecular Biology Laboratory). On its own, *v-erb-A* is not able to transform cells. Instead, in the avian erythroblastosis virus, it cooperates with the *v-erb-B* oncogene, which induces the self-renewal of erythroblasts without completely blocking their differentiation into mature erythroid cells. The block is completed by the product of *c-erb-A*, which perhaps achieves this by blocking the nuclear transcription of the erythroid membrane protein band III (H. Beug, European Molecular Biology Laboratory).

The struggle to understand the function of the *src* and *ras* proto-oncogenes and oncogenes continues. For some time it has been clear that the *c-src* product in neural tissues is a variant of the protein found elsewhere. It is now clear that the variation is the result of a 6-amino-acid inser-

tion between the sequences encoded by exons three and four (J. Brugge, State University of New York, Stonybrook; A. Bernards, Whitehead Institute, Boston). The *src* product is a protein tyrosine kinase and the latest of its substrates to be in the limelight had seemed to be related to the glucocorticoid hormone-induced inhibitors of phospholipase A. But further scrutiny of these substrates, and particularly p36, suggests that they are unrelated to the inhibitors and instead are membrane proteins that use their phospholipid-binding properties to link the plasma membrane to the cytoskeleton of cells (T. Hunter, Salk Institute).

For *ras* proteins, there is continued speculation that they serve as mediators of growth-factor induction of inositol phospholipid turnover. Recent evidence that is compatible with this popular notion has been extended to suggest that the products of the three *ras* proto-oncogenes preferentially couple different growth-factor signals to inositol phospholipid turnover and that constitutive turnover is conferred by the presence of *ras* oncogenes

(C. Marshall, Institute of Cancer Research, London).

The fact that in another cell system *ras* has no effect on phospholipid turnover (J. Pouyssegur, CNRS Nice) serves as a timely reminder that not all cells are the same. The additional hazard of much oncogene research is that it is heavily dependent on particular cell lines and gene constructs. Thus there is always room for some concern about its relevance to physiological systems. It was encouraging, therefore, to learn that the ability of *ras* and *myc* to cooperate in the transformation of cultured cells can be adapted to a reconstituted embryonal prostate system, where this cooperation results in tumours (H. Land, Imperial Cancer Research Fund, London). The meeting demonstrated the healthy state of research into oncogenes and growth control which has certainly grown into the kind of mini-industry that justifies a new journal (or two?) and the decision to hold a similar meeting again in 18 months' time. □

Peter Newmark is Deputy Editor of Nature.

Solar System

Pluto, Charon and eclipses

David W. Hughes

PLUTO, the outermost known member of the Solar System, is a curiosity both for its eccentric orbit and for its diminutive size and mass, setting it apart from the other planets. The sizes of Pluto and its satellite Charon have recently been measured with considerable precision, thanks to a rare series of mutual eclipses and occultations. Computer modelling has been used to fit the variations in the amount of scattered light as a function of time. Typical results, obtained by M. Pakull and K. Reinsch¹, indicate that the radii are $1,100 \pm 70$ km (Pluto) and 580 ± 50 km (Charon). On page 127 of this issue², Tedesco and colleagues report similar values, $1,100 \pm 75$ and 650 ± 75 km, respectively. So Pluto is very small for a planet, 63 per cent of the size of our Moon. Also, Pluto and Charon are unusual, having the highest mass ratio of any pair of gravitationally locked Solar System objects.

Pluto turns out to be much less massive than was once believed. In the 1930s, just after its first sighting, the mass was estimated to be about the same as that of Earth. The discovery of Charon in 1978, placed the mass measurement on a firmer footing. Kepler's third law gives the mass of Pluto, M_p , and Charon, M_c , simply as

$$M_p + M_c = 4\pi^2 a^3 / G P^2$$

where a is the mean distance between Pluto and Charon, G is the gravitational constant and P is the orbital period of the satellite around Pluto. In late 1978 Pluto's

mass was provisionally calculated to be about 1/380 the mass of Earth.

The system has become tidally locked. Charon orbits Pluto every 6.38718 days and this is the same (to at least ± 0.0003 per cent) as the spin periods of both. One face of Pluto always faces one face of Charon, and Charon's orbit is circular and in Pluto's equatorial plane. Between the 1950s and the present the brightness of Pluto has decreased by about 32 per cent. Also the amplitude of the periodic variation of brightness has become larger. Pluto, like Uranus, has a spin axis which is highly inclined with respect to the normal of the orbital plane. In the 1950s we were looking along Pluto's axis at one of the bright, uniform, polar regions. Nowadays the view is more equatorial and Pluto's tropics seem to be made up of a series of bright and dark areas, so that the amount of light scattered varies considerably as it spins.

At present the Earth is very close to the orbital plane of Charon and astronomers can see a series of eclipses. This edge-on geometry occurs twice during each 248 years of Pluto's orbit of the Sun. The eclipse season lasts about 5 years, the present one ending in about 1990. A table of the 1987 eclipses has recently been produced³ by David J. Tholen and colleagues. Superior events occur when Pluto passes in front of Charon and the magnitude changes by about 0.24. The eclipse lasts for

between 32 and 79 minutes, depending on how nearly the centres of the two objects line up. Inferior events occur when Charon crosses the face of Pluto. These are not always total because Charon and its shadow are usually not encompassed by Pluto's projected disk. The typical magnitude change is about 0.55. Superior and inferior eclipses are seen alternately every 3.19 days, the season lasting from mid-December to mid-September, while Pluto is well separated from the Sun in the sky.

Eclipse observations are usually obtained with large telescopes, for example, the 2.2-m reflector at the European Southern Observatory, La Silla, Chile; the 2.24-m telescope of the Mauna Kea Observatory, Hawaii; and the 2.1-m telescope of the McDonald Observatory, Texas. Standard blue and green filters are used and one photometric measurement is taken about every 4 minutes. In the recent report of Pakull and Reinsch¹, light curves (of the temporal variation) are analysed using models that have been generated for eclipsing binary stars. It is assumed that the two components, Pluto and Charon, are spherical and that they are uniformly bright. The best fit indicates that the radii of Pluto and Charon are $1,100 \pm 70$ km and 580 ± 50 km, respectively. These values can be compared with those obtained by Dunbar and Tedesco⁴, $1,150 \pm 50$ km and 750 ± 50 km respectively, and D.J. Tholen and colleagues (personal communication), $1,145 \pm 46$ km and 642 ± 34 km respectively. A different approach to the measurement is reported in this issue², where Tedesco *et al.* used a single 60- μ m wavelength observation of Pluto and Charon made with the Infrared Astronomical Satellite (IRAS) in July 1983. The measured flux was 0.42 ± 0.04 Jy and this, coupled with an assumed emissivity, gives radii of $1,100 \pm 75$ km and 650 ± 75 km, respectively.

D.J. Tholen⁵ obtained a more precise value for the radius of Charon's orbit around Pluto, so that a in the above equation is $19,130 \pm 460$ km giving the total mass of the system as 1.36×10^{25} g (1/439 Earth masses). The mean density is thus 1.84 ± 0.19 g cm⁻³ and the volume of Pluto is 5.7 times that of Charon. Unfortunately the individual masses and densities are not known but the system must be a mixture of rock and ice, with rock predominating. The mean density is very close to that of Ganymede, Callisto, Titan and Triton, four of the larger moons in the outer Solar System.

Pluto has a slightly reddish tinge and a high visual reflectivity of 61 per cent. The reflectivity of Charon is 42 per cent, the side facing away from Pluto having a similar colour to that of Pluto but the side pointing towards Pluto being more neutral. These differences in reflectivity indicate that Charon is warmer than Pluto. The present model of Pluto's surface

features⁶ has bright polar caps and a less reflective equatorial band, the band sporting two large spots, one bright and one dark. The bright regions are possibly exposed, bare, solid ice.

So Pluto is really an insignificant object; even our Moon is 6.4 times more massive. It is certainly not planetary, and its mass and density point to it being an escaped satellite. But which planet did it escape from, and when? And why did the perturbations responsible for the escape not rip the Pluto-Charon system apart? I still think that Neptune is the probable parent, and it is worth noting that the Pluto-Charon mass ratio (assuming that Pluto has a density slightly larger than Charon) is close to the values obtained by Sir James Jeans⁷ when he studied the break up of

condensing, rapidly rotating fluid bodies. Maybe Charon was born in a similar fashion only after Pluto had escaped from Neptune's family of satellites, all this happening in the early stages of planetary formation. □

1. Pakull M. & Reinsch K. *Europ. Space Observ. Messenger* No. 46, 3-6 (1986).
2. Tedesco, E.F., Veeder, G.J. & Dunbar, R.S. *Nature* **327**, 127-129 (1987).
3. Tholen, D.J., Buie, M.W. & Swift, C.E. *Astr. J.* **93**, 244-247 (1987).
4. Dunbar, R.S. & Tedesco, E.F. *Astr. J.* **92**, 1201-1210 (1986).
5. Tholen, D.J. *Astr. J.* **90**, 2353-2359 (1985).
6. Buie, M.W. & Tholen, D.J. *Bull. Am. Astr. Soc.* **18**, 821 (1986).
7. Jeans, J.H. *Astronomy and Cosmogony*, Ch.8 (Cambridge University Press, 1928).

David W. Hughes is Senior Lecturer in Astronomy and Physics at the University of Sheffield, Sheffield S3 7RH, UK.

Cytoskeleton

Tubulin's terminal tyrosine

Roy Burns

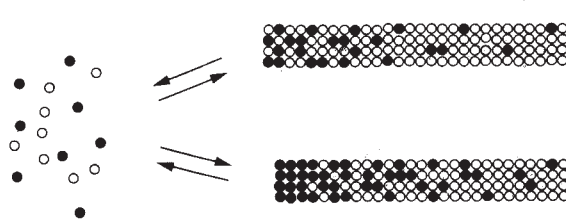
IT IS now more than 10 years since a novel post-translational modification of the protein tubulin, the principal structural component of microtubules, was first identified. The modification involves the addition of tyrosine to the carboxy terminus of α -tubulin, catalysed by a tubulinyl tyrosine ligase (TTL) with the concomitant hydrolysis of ATP—that is, no messenger RNA or ribosomes are involved¹. The use of tyrosine is not totally specific: phenylalanine and DOPA are also substrates for TTL. A specific tubulinyl tyrosine carboxypeptidase (TTC) has also been purified², which removes this terminal tyrosine to yield α -tubulin terminating in a glutamic acid residue (that is, Tyr-Tu subunits are converted into Glu-Tu). These two enzymes offer the potential of a cycle of tyrosine addition and loss, and clues are now emerging about the role of this unusual cycle (see figure).

First, the coding portion of every gene encoding α -tubulin yet sequenced (with an exception which I will discuss later) terminates in a tyrosine codon, indicating that the primary gene product is tyrosinylated (Tyr-Tu). Indeed, Tyr-Tu is very widespread, as a monoclonal antibody (YL1/2)³ raised against yeast tubulin specifically recognizes the carboxy-terminal sequence Gly-(Glu)₃-Gly-(Glu)₂-Tyr (ref. 4). This antibody, made by John Kilmartin (Laboratory of Molecular Biology, Cambridge), has been widely used to examine the microtubule cytoskeleton of various animal, plant and lower eukaryotic cells, with the result that the epitope (antigenic determinant) is now known to be highly conserved.

The second clue comes from studies of the substrate specificity of TTL and TTC, with TTL specifically using the unassembled tubulin dimers⁵ whereas

Microtubules, a major structural component of the cytoskeleton of all eukaryotic cells, are assembled from $\alpha\beta$ tubulin dimers. The open symbols represent Glu- $\alpha\beta$ dimers and the closed symbols Tyr- $\alpha\beta$ dimers. The deetyrosinylation by TTC, acting on the assembled microtubule,

leads to a steadily increasing proportion of Glu- $\alpha\beta$ dimers with time, and hence with distance along the microtubule, such that 'old' microtubules contain only Glu- $\alpha\beta$ subunits. In the upper model microtubule, both Glu- and Tyr- $\alpha\beta$ dimers are added to the microtubule end. The deetyrosinylation may signal the binding of microtubule-associated proteins or other molecules (not shown) that stabilize the microtubule, leading to the formation of a subpopulation of 'old', and hence Glu- $\alpha\beta$ tubulin, microtubules. In the lower model, only Tyr- $\alpha\beta$ dimers are added to the microtubule end, so that the newly assembled microtubule consists solely of tyrosinylated α -tubulin. The proportion of Tyr- α :Glu- α , determined by the rate of *de novo* protein synthesis and the activity of the TTL, therefore controls the size of the assembly-competent pool size. The figure shows the gradient of the increasing proportion of Glu- $\alpha\beta$ subunits over a short length of the microtubule, which *in vivo* is predicted to take considerable time and hence would be spread over several thousand subunits.



TTC preferentially acts on assembled microtubules⁶. This result suggests that tubulin tyrosinylation is involved in the assembly status of the tubulin. To unravel the role of tyrosinylation, Gundersen and Bulinski⁷ raised polyclonal antibodies that specifically recognize Tyr-Tu and Glu-Tu, and examined the distribution of Tyr-Tu and Glu-Tu microtubules in several tissue-culture cells, in dividing and non-dividing cells, and in specific cellular structures containing microtubules^{7,9}. Their findings suggest that microtubules that are rapidly turning over tend to have a high Tyr-Tu content whereas more stable microtubules predominantly contain Glu-Tu. Similarly, Schultz and Kirschner¹⁰, who performed blocking experiments with these polyclonal antibodies, showed that stable microtubules consist of Glu-Tu whereas newly assembled microtubules react with the antibody against Tyr-Tu.

This relationship between tyrosinylation and the age of assembled microtubules is clearly shown in studies on the intracellular distribution of Tyr-Tu during the cell cycle of *Trypanosoma brucei brucei*¹¹. These trypanosomes have a prominent flagellum and a sub-pellicular cytoskeleton, both formed of stable microtubules which are replicated during cell division. As the cell divides, its posterior part (where the sub-pellicular microtubules are believed to be assembled) and the newly formed daughter flagellum react strongly with the YL1/2 monoclonal antibody, but this immunoreactivity is gradually lost as replication of the cytoskeleton and the flagellum is completed. Consequently, the newly assembled microtubules contain a high content of Tyr-Tu subunits which are subsequently converted to Glu-Tu, showing clearly that these stable microtubules are not derived from the exclusive assembly of the Glu-Tu subunits of the cytoplasmic pool.

There are two alternative explanations for these observations. First, the stable sub-population of microtubules may be stable *because* they contain Glu-Tu: the loss of the tyrosine residue may specify the binding of microtubule-associated proteins or other molecules that stabilize microtubules. Alternatively, all newly assembled microtubules may only be formed of Tyr-Tu, with the increasing Glu-Tu content simply reflecting the longer time available for de-tyrosinylation by TTC. This raises the interesting prospect that *in vivo* only the Tyr-Tu is assembly-competent. Such subunits would be generated either by *de novo* synthesis, as the primary gene product is tyrosinylated, or as a result of the charging of Glu-Tu by TTL. The newly assembled Tyr-Tu microtubules would then become substrates for TTC, so that the more stable microtubules would have an increasingly high Glu-Tu content. The subunits returned to the cytoplasmic pool

as a result of microtubule disassembly would, depending on the age of the microtubules and the activity of the TTC, progressively have a higher proportion of Glu-Tu subunits and would need to be retyrosinylated by TTL before being reassembled. In such a model (see figure), the function of tyrosinylation would be to regulate the size of the assembly-competent tubulin pool, similar to the role of profilin in regulating assembly of actin into filaments.

Unfortunately, there is no direct evidence for the selective assembly of Tyr-Tu from *in vitro* assembly studies using conventional reassembly buffers^{5,12}, but this could indicate more about the inadequacies of the studies of microtubule assembly *in vitro* than about the *in vivo* control mechanism. The next step in understanding the cellular function of tyrosinylation will surely focus on whether there is selective use of Tyr-Tu from the cytoplasmic pool. Such studies will undoubtedly be difficult as immunofluorescent techniques are far more powerful when considering the assembled microtubules than they are for examining the composition of the cytoplasmic pool.

Just when it seemed that a coherent story was emerging about the role of tyrosinylation, Monteiro and Cox report¹³ data that hint at a whole new level of complication. The inferred amino-acid sequence of a gene encoding α -tubulin from the slime mould *Physarum polycephalum* has a carboxy-terminal methionine rather than

the ubiquitous tyrosine, yet other *Physarum* α -tubulins crossreact with the YL1/2 antibody. Although the preceding terminal amino acids are identical to those of more conventional α -tubulins, the presence of the terminal methionine strongly suggests that it is not a substrate for TTC, and hence could not be converted into a form capable of being tyrosinylated by TTL. The cloned gene may be encoding the plasmodia-specific α -tubulin isotype that is not recognized by the YL1/2 antibody. It is unclear whether this novel *Physarum* α -tubulin is telling us something new about tyrosinylation or whether it is the exception that proves the rule. □

1. Arce, C. A., Rodriguez, J. A., Barra, H. S. & Caputto, R. *Eur. J. Biochem.* **59**, 145-149 (1975).
2. Argarana, C. E., Barra, H. S. & Caputto, R. *J. Neurochem.* **34**, 114-118 (1980).
3. Kilmartin, J. V., Wright, B. & Milstein, C. *J. Cell Biol.* **93**, 576-582 (1982).
4. Wehland, J., Willingham, M. C. & Sandoval, I. V. *J. Cell Biol.* **97**, 1467-1475 (1983).
5. Arce, C. A., Hallak, M. E., Rodriguez, J. A., Barra, H. S. & Caputto, R. *J. Neurochem.* **31**, 205-210 (1978).
6. Kumar, N. & Flavin, M. *J. biol. Chem.* **256**, 7678-7686 (1981).
7. Gundersen, G. C., Kalnoski, M. H. & Bulinski, J. C. *Cell* **38**, 779-789 (1984).
8. Gundersen, G. C. & Bulinski, J. C. *J. Cell Biol.* **102**, 1118-1126 (1986).
9. Gundersen, G. C. & Bulinski, J. C. *Eur. J. Cell Biol.* **42**, 288-294 (1986).
10. Schultz, E. S. & Kirschner, M. *J. Cell Biol.* **103**, 432a (1986).
11. Sherwin, T., Scheider, A., Sasse, R., Seebeck, T. & Gull, K. *J. Cell Biol.* **104**, 439-446 (1987).
12. Kumar, N. & Flavin, M. *Eur. J. Biochem.* **128**, 215-222 (1982).
13. Monteiro, M. J. & Cox, R. A. *J. molec. Biol.* **193**, 427-438. (1987).

Roy Burns is a Lecturer in the Biophysics Section, Blackett Laboratory, Imperial College of Science and Technology, London SW7 2BZ, UK.

Neuropharmacology

Histamine receptors branch out

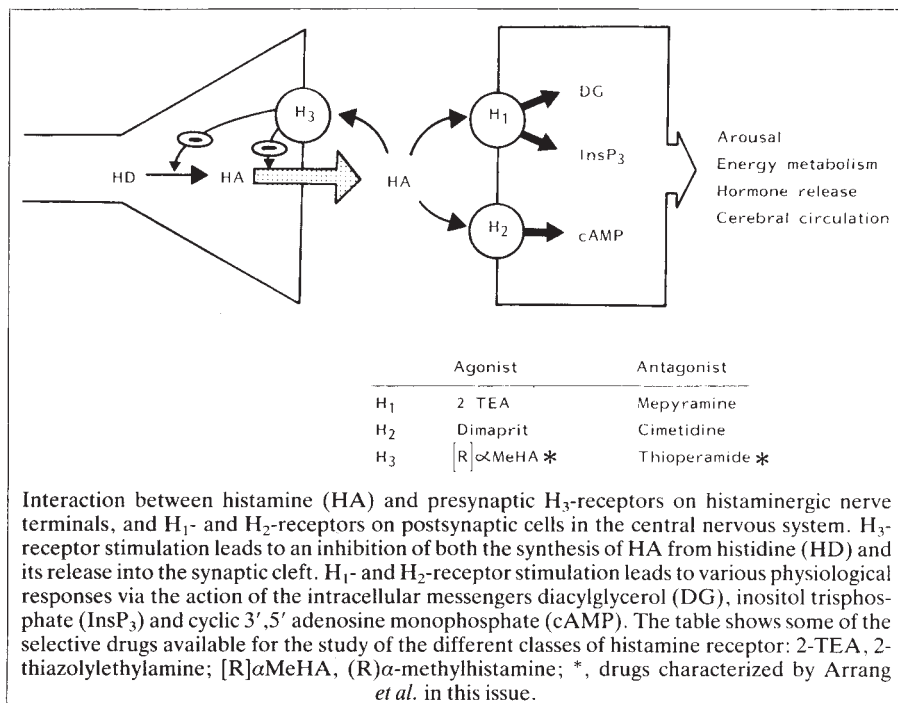
Stephen J. Hill

HISTAMINE has long been known to be distributed widely in the mammalian central nervous system, but it is only recently that neuronal pathways containing histamine and its synthesizing enzyme histidine decarboxylase have been seen using immunohistological techniques¹. Histamine receptors have been classified into two main subtypes, H₁ and H₂, on the basis of quantitative studies in isolated peripheral tissues^{2,3}. Both types of receptor exist in brain tissue and are linked to different intracellular messenger systems through which they may mediate various physiological functions (see figure)^{4,5}. The article by Arrang *et al.* on page 117 of this issue⁶ provides strong evidence for the presence of a novel histamine receptor (H₃) on histaminergic nerve endings in brain tissues controlling the synthesis and release of histamine (see figure). Furthermore, Ishikawa and Sperelakis show on page 158 of this issue⁷ that H₂-receptors exist outside the brain and that their location is not confined

to histamine-containing nerve endings.

The groundwork for the new studies was laid by Arrang *et al.*⁸ who reported in 1983 that histamine could inhibit its own release from rat cerebral cortex, and established that the characteristics of the receptor controlling the release of histamine were different from those of the two known classes of histamine receptor. But confirmation of the existence of a new class (H₃) of histamine receptor required agents that either selectively stimulated (agonists) or selectively inhibited (antagonists) this response. Arrang *et al.* in this issue⁶ now describe the properties of two selective, potent ligands, thioperamide and (R) α -methylhistamine, for this novel H₃-receptor.

Thioperamide, a competitive antagonist of the H₃-receptor in brain slices, is effective in the nanomolar concentration range and has a negligible effect on H₁- and H₂-receptor-mediated responses. Before this study the only potent antagonists of the H₃-receptor were impromidine



(also a potent H₂-agonist), betahistine (H₁-agonist) and burimamide (H₂-antagonist), which lack the necessary specificity of action. Arrang *et al.*⁶ also developed (R)α-methylhistamine as a potent and selective H₃-receptor agonist by capitalizing on the marked stereoselective agonist requirements of this receptor. We are now in a stronger position to study the H₃-receptor than the classical H₁-receptor, which has no agonists of similar high selectivity. Radiolabelling (R)α-methylhistamine allows H₃-receptors to be seen in different regions of rat brain and in guinea-pig lung tissue.

It seems likely that H₃-receptors in peripheral tissues have a similar role to those in the brain and control the synthesis and release of histamine. Arrang *et al.*⁶ show that stimulation of H₃-receptors with (R)α-methylhistamine can reduce histamine formation in the lung, spleen and skin of the rat. They suggest that these H₃-receptors are present on mast cells, which are abundant in these tissues, and control the synthesis and release of histamine and, perhaps, other mediators of allergy and inflammation. Note, however, that histamine synthesis can occur in other cell types in these tissues and can be markedly induced during rapid tissue growth and regeneration (for example, skin cells following injury and during tumour development⁹). H₃-receptors may also be present on these cells.

The existence of H₃-receptors outside the brain is also suggested in this issue by Ishikawa and Sperelakis⁷, who present evidence for a role for H₃ receptors in the control of the microcirculation. They show that histamine can depress the release of noradrenaline, the chemical transmitter mediating the contraction of

vascular smooth muscle cells produced by sympathetic nerve stimulation. The selective ligands described by Arrang *et al.*⁶ will confirm whether the response is mediated by H₃-receptors.

One can predict that the availability of (R)α-methylhistamine and thioperamide will trigger a marked increase in interest in histamine research similar to that following the report by Black and colleagues of the properties of the first H₂-antagonist, burimamide³. Indeed, because of the marked antagonist potency of burimamide at the H₃-receptor, it may be profitable to re-evaluate those 'H₂-

responses' that were originally classified with burimamide but not with the other more specific H₂-antagonists that later became available. Thioperamide may also be useful in establishing the function of histamine in the central nervous system as it can cross the blood-brain barrier and increase histaminergic neuron activity *in vivo* by removing the normal feedback inhibitory control on histamine release and synthesis (see figure). The data of Ishikawa and Sperelakis⁷, however, suggest that H₁-receptors are not exclusively located on histamine nerve terminals and may control the release of other neurotransmitters. Thus, it is important to determine whether administration of thioperamide can increase the release of other chemical messengers in the central nervous system in addition to histamine.

It remains to be seen whether the new H₃-selective ligands or their descendants will follow the selective H₁- and H₂-antagonists into clinical use, but there is therapeutic potential in cardiovascular regulation, allergy and mental illness. □

1. Pollard, H. & Schwartz, J.-C. *Trends Neurosci.* **10**, 86-89 (1987).
2. Ash, A.S.F. & Schild, H.O. *Br. J. Pharmac. Chemother.* **27**, 427-439 (1966).
3. Black, J.W., Duncan, W.A.M., Durant, C.J., Ganellin, C.R. & Parsons, M.E. *Nature* **236**, 385-390 (1972).
4. Prell, G.D. & Green, J.P. *Ann. Rev. Neurosci.* **9**, 209-254 (1986).
5. Schwartz, J.-C., Arrang, J.-M. & Garbarg, M. *Trends pharmac. Sci.* **8**, 24-28 (1986).
6. Arrang, J.-M. *et al.* *Nature* **327**, 117-124 (1987).
7. Ishikawa, S. & Sperelakis, N. *Nature* **327**, 158-160 (1987).
8. Arrang, J.-M., Garbarg, M. & Schwartz, J.-C. *Nature* **302**, 832-837 (1983).
9. Bartholeyns, J. & Fozard, J.R. *Trends pharmac. Sci.* **7**, 123-125 (1985).

Stephen J. Hill is in the Department of Physiology and Pharmacology, Queen's Medical Centre, University of Nottingham Medical School, Nottingham NG7 2UH, UK.

Quantum mechanics

Gases exist — official!

Ian Stewart

THE title of this article may raise eyebrows. No experimentalist would either deny it or think it newsworthy. But theoreticians will consider it an overstatement — no one has yet been able to deduce, by rigorous mathematical analysis, that the fundamental laws of quantum mechanics imply the existence of ordinary states of matter such as crystals or gases. Now that goal no longer looks unattainable, as Charles Fefferman of Princeton University explains in a recent paper (*Commun. pure appl. Math.* **39**, S67-S109; 1986). Indeed, the admittedly partial results already obtained come impressively close, however difficult it may be to bridge the final gaps. The paper is part of a special supplement devoted to the symposium *Frontiers of the Mathematical Sciences* held to mark the fiftieth

anniversary of the Courant Institute: the entire volume makes fascinating reading for anyone interested in the relation between mathematics and its applications.

In earlier times it was common to talk of the laws of nature and to assume that their mathematical formulation was an exact statement of the behaviour of the physical world. Today it is generally held that the equations of physics are more appropriately viewed as simplified models, valid within their own sphere, but lacking the authority of ultimate truth. Having been forced to revise 'ultimate truth' more than once, scientists are reasonably enough playing safe.

Some models, however, are more fundamental than others; and at the level of small-scale structure, quantum mechanics reigns supreme. It says a great deal

about the behaviour of elementary particles, such as electrons and protons, or of small systems of particles such as the hydrogen atom, a proton-electron pair. It is also, especially when subjected to modern revisions that take relativistic effects into account, a theory whose accuracy is breathtaking. Agreement with experiment to eight decimal places is not unusual.

Quantum mechanics is less successful for systems containing many particles. An engineer contemplating the buckling of an iron strut under compressive forces does not rely greatly on quantum mechanics. Instead he applies either a classical continuum model of elastic buckling, or a finite element model using a computer, and supplies his model with basic data such as the elastic constants of an iron bar.

Most quantum theorists believe, as an article of faith, that the large-scale behaviour of an iron bar is a logical consequence of the quantum mechanics of its constituent particles. There are many arguments, experiments and approximate calculations to underpin this belief, but it has never been established by rigorous mathematics. Nobody has ever computed elastic constants from a quantum-mechanical model of an iron bar treated as an assembly of nuclei and electrons. The main obstacle is the notorious difficulty of studying systems of many bodies in terms of two-body interactions. Fefferman's paper deals with the quantum mechanics of a system of electrons and nuclei interacting by inverse-square forces. Its aim is to prove, using only quantum mechanics, that such a system can exist in a state that corresponds to an atomic gas.

In quantum mechanics the state of a system is represented by a complex-valued wavefunction Ψ . The outcome of any experiment performed on the system is determined by Ψ . In particular the absolute value $|\Psi|^2$ can be interpreted as the probability of finding the particles in given positions.

Naturally, Ψ is hard to obtain. Much easier is the hamiltonian H , which can be written down explicitly in terms of the inverse-square law for particle interactions. The hamiltonian can be described roughly as the quantum-mechanical analogue of classical energy, but it is a linear operator on a Hilbert space rather than just a number. Associated with it is a series of numbers, its eigenvalues, $E_1, E_2, \dots$ which represent the energies of the pure quantum states $\Psi_1, \Psi_2, \dots$ corresponding to H .

For systems with only a few particles it is sensible to find a Ψ that minimizes the average energy of the system. Consider, for example, a model hydrogen atom: a proton at the origin plus a single quantized electron. The ground-state energy (lowest eigenvalue of H) is -0.25 , and the wavefunction that minimizes the average

energy is $\Psi(x) = (1/\sqrt{8\pi})\exp(-1/2|x|)$. This mathematically derived wavefunction makes good physical sense.

For systems with many particles, however, it is not sensible to insist that Ψ minimizes the energy, as that describes matter at absolute zero temperature, which is not the case of prime interest. Instead, at non-zero temperature T , the wavefunction Ψ is defined by thermodynamic principles. The particles are confined in a box, and the wavefunction is defined to be that corresponding to a pure quantum state Ψ_k , picked at random according to a probability proportional to $\exp(-E_k/T)$. To make the sum of probabilities 1, these numbers must be divided by the partition function $Z = \exp(-E_1/T) + \exp(-E_2/T) + \dots$. This formalism provides a complete statistical description of the particles in the box. In particular it determines the Gibbs measure, the probability of the particles being in given positions.

There are now two fundamental questions. The first, basic to statistical mechanics, is: how does the Gibbs measure behave as the box becomes infinitely large? This is the mathematical version of the question: what does ordinary matter look like at a given temperature and density? The question is interesting because the answer — crystal, quasicrystal, liquid, gas or whatever — depends on temperature and density in a way that is far from understood, even empirically.

The second question is basic to thermodynamics: what is the behaviour of the limit F of $\log Z$ divided by the volume of the box? This is important because all major thermodynamic quantities — pressure, entropy, specific heat and so on — can easily be derived from F . There is as yet no chance of computing F explicitly. In fact it is a deep theorem that the limit F exists (Lebowitz, J. & Lieb, E. *Adv. Math.* **9**, 316–398; 1987).

Fefferman has studied the first question, attempting to prove that for a suitable range of temperature and density the Gibbs measure describes an atomic gas. In such a gas, the particles come in electron-proton pairs, these pairs being widely separated in comparison to their own size. The goal is to prove the theorem that there are values of temperature and density such that, in a large enough box, the Gibbs measure describes a gas of atoms in which all but a vanishingly small proportion of the atoms are as widely separated as we please.

This theorem has been proved, except for one technical assumption. This assumption, the crucial gap in formulating a rigorous argument, is thus of great interest. Ironically, the analogous experimental fact is easily established. It concerns an estimate for the binding energy E_* of bulk matter, defined as the

energy per atom needed to disperse the matter to infinity while imparting zero velocity. The experimental value (found by reversing the procedure and freezing hydrogen to form a crystal) is ~ 0.3 . For a single atom the binding energy is less, 0.25.

The assumption needed to prove the theorem is (roughly) that the binding energy E_* for bulk matter is no more than 0.5. But this is far from easy to establish as a rigorous mathematical consequence of quantum mechanics. In fact, the finiteness of the binding energy was first proved by F. Dyson and A. Lenard (*J. Math. Phys.* **8**, 423–434; 1967 and **9**, 698–711; 1968). Their estimate, $E_* \sim 10^{14}$, is a far cry from what is needed here. It was later improved by E. Lieb and W. Thirring (*Ann. Phys.* **155**, 494–512; 1984), and by E. Lieb (*Am. Math. Soc. Proc. Symp. pure Math.* **36**, 241–252; 1980) to 5.4. Attempts to improve this estimate have focused on a particular mathematical conjecture made by Lieb and Thirring.

Fefferman had the idea that the Lieb-Thirring conjecture might actually be harder than the calculation of E_* , so he developed a new method with the assistance of Hale Trotter. The method has not yet worked for electrons and protons (hydrogen atoms) but it has succeeded for electrons and nuclei of charge 40–90, which includes xenon atoms. This method leads to a 'proof' in the sense that E_* is estimated rigorously in terms of the eigenvalues of an explicit hamiltonian acting on an explicit — and simple — Hilbert space. These eigenvalues are then computed numerically. To complete the proof, it is sufficient to reorganize the computer calculation so that it provides not just a numerical answer (subject to unknown truncation errors) but a rigorous bound. This is an interesting question in its own right, and it is cropping up in many problems on the borderline between mathematics and applied science.

Fefferman's paper contains several other interesting results, notably on the relativistic stability of hydrogen. It also describes what further work needs to be done. The first goal is to complete the proof of the main theorem on the existence of a gas of hydrogen atoms, by consolidating the computer calculations, by improving the estimate of E_* and by reducing the nuclear charge from 40 to 1. Beyond lie even more difficult questions — for example, why does matter arrange itself in crystalline lattices at low temperatures? There remains much to understand about the implications of quantum mechanics for bulk matter; but the prospect of making rigorous connections between the two, not so long ago negligible, is beginning to look rather good. $\square$

Ian Stewart is at the Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK.

Origins of HTLV-4

SIR—The recent report¹ of the remarkable genetic similarity of human T-lymphotropic virus type-4 (HTLV-4) with certain simian immunodeficiency virus (SIV) isolates from both macaques and African green monkeys (AGMs) and the accompanying editorial² discussing the possibility of laboratory-acquired contamination have prompted us to outline a chronology of events as they bear on this discussion.

SIV was first isolated from macaques by us at the New England Regional Primate Research Center³. The first three isolates, termed 251, 220 and 239, were obtained in September and October 1984. Once it was confirmed by electron microscopy at the centre that these T-cell tropic virus isolates were retroviruses with lentivirus morphology (like HIV), we gave SIV-producing HUT-78 cell cultures to Drs Kanki and Essex at the Harvard School of Public Health in November–December 1984 for antigenic protein analysis. The 251 isolate was the prototype primarily used in these early studies⁴. The description of SIV-reacting antibodies in AGMs⁵ and humans in West Africa⁶ followed the initial isolation and characterization of macaque-SIV and led to reports of isolation of immunodeficiency viruses from AGMs⁷ and West Africans⁸ by the Essex laboratory. At the time the macaque-SIV producing cultures were given to Kanki and Essex, they were the only cultures in our laboratory producing monkey or human AIDS virus and were unquestionably derived from infected macaques.

The limited strain variability (0–3 restriction site polymorphisms out of 20 sites) in the four macaque-SIV isolates from our centre that we have examined to date (251, 239, 157 and 142) is not surprising. Certainly 251 and 239 should be similar as animal 239 was inoculated with infected material originating from the 251 animal³. Furthermore, SIV from the 157 animal was transmitted *in utero* to her offspring 142. All of our isolates have been obtained from a closed colony of macaques where infection is quite rare; only 3 of 848 macaques in our 1986 survey were SIV seropositive. It thus seems likely that a single strain of SIV introduced into our colony eight or more years ago has spread to different animals in the colony.

But remarkably isolates obtained from Central African green monkeys⁷ and West African humans⁸ in a single laboratory at the Harvard School of Public Health turn out to be 99 per cent homologous with the macaque-SIV isolate being carried in the same laboratory. Our analyses, similar to those reported by Kornfeld *et al.*¹, indicate that the macaque SIV prototype strain 251 is identical at greater than 20 restriction endonuclease sites with the maps now published for SIV-AGM⁸ and HTLV-4¹. It

has been suggested that resolution of the problem regarding the authenticity of the SIV-AGM and HTLV-4 isolates “will probably require the isolation under stringent conditions of further examples of the viruses”². It appears, however, that a considerable number of such isolates have already been compared. Kornfeld *et al.*¹ report that six AGM isolates and two HTLV-4 isolates from the Essex laboratory “were identical at 32 sites”; SIV isolates from a sooty mangabey and from a macaque at the California and Delta Regional Primate Research Centers showed considerable variation between each other and with the isolates in question. We have analysed a mangabey-SIV obtained by P. Fultz at yet another primate research centre⁹ that had been passed in our laboratory for over a year and it too showed considerable differences. HIV-2 isolates obtained by the French from humans in West Africa have also shown considerable variation when compared among themselves and with the other isolates in question¹⁰.

R.C. DESROSIERS

M.D. DANIEL

N.L. LETVIN

N.W. KING

R.D. HUNT

*New England Regional Primate Research Center,
Harvard Medical School,
One Pine Hill Drive,
Southborough, Massachusetts 01772, USA*

1. Kornfeld, H. *et al. Nature* **326**, 610–613 (1987).
2. Newmark, P. *Nature* **326**, 548 (1987).
3. Daniel, M.D. *et al. Science* **228**, 1201–1204 (1985).
4. Kanki, P. *et al. Science* **228**, 1119–1201 (1985).
5. Kanki, P. *et al. Lancet* **i**, 1330–1332 (1985).
6. Kanki, P. *et al. Science* **232**, 238–243 (1986).
7. Kanki, P. *et al. Science* **230**, 951–954 (1985).
8. Hirsch, V. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 9754–9758 (1986).
9. Fultz, P. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 5286–5290 (1986).
10. Clavel, F. *et al. Nature* **324**, 691–695 (1986).

Is aluminium leaching enhanced by fluoride?

SIR—Tennakone and Wickramanayake¹ have described a 1,000-fold enhancement of aluminium leaching from cooking utensils when 1 p.p.m. fluoride is present in the cooking water. This observation is of considerable concern to both pro-

ponents and opponents of water fluoridation, especially because the question of neurotoxicity of ingested aluminium in individuals with normal renal function is unresolved. The data on the amount of aluminium leached are in fact so surprising that we have repeated the experiments described by Tennakone and Wickramanayake¹. We obtained 15 aluminium pans of different sizes, 12 of which had experienced considerable use and 3 were new. All were cleaned scrupulously before use in the leaching experiments.

We followed the experimental protocol of Tennakone and Wickramanayake¹. We boiled 450 ml of a solution of citric acid in reverse-osmosis (RO)-treated water (2.0 mM, pH 3.10) for 10 min and repeated the study using the same solution containing sodium fluoride (0.024 mM, 1 p.p.m.). We continued to follow the original protocol¹ by boiling crushed fresh tomatoes (50 g in 250 ml of RO water) also in the presence and absence of sodium fluoride (0.024 mM). All solutions were analysed for aluminium content using electrothermal atomic absorption spectrometry with Zeeman background correction using a Perkin-Elmer model 5100 instrument. The method was that described earlier by us² and used aqueous aluminium salt solutions prepared from material certified by the National Bureau of Standards (SRM 2127-1) for calibration. The validity of the analytical method has been verified by recovery, linearity and interference studies, as well as by interlaboratory proficiency test surveys.

The result of the studies are given in the table. The citric acid solutions, before they were boiled in the aluminium pans, had an aluminium concentration of 0.081 mM without fluoride and 0.107 mM with fluoride. The solutions containing the crushed fresh tomatoes had aluminium concentrations of 0.355 mM and 0.215 mM without and with 0.024 mM sodium fluoride, respectively. The pH of these crushed tomatoes in water ranged from 4.43 to 4.49.

Our studies demonstrate minimal enhancement by fluoride of aluminium leaching from aluminium cooking utensils. We postulate three possible reasons for the discrepancy between our findings and those of Tennakone and Wickram-

Aluminium leaching data

	Aluminium leached without fluoride (mg)	Aluminium leached with fluoride (mg)	Ratio of aluminium leached + NaF/– NaF
	Mean ± s.d.	Mean ± s.d.	Mean ± s.d.
Citric acid solution	7.08 ± 5.11	8.19 ± 5.51	1.29 ± 0.53
Crushed tomatoes	0.98 ± 0.58	0.69 ± 0.41	0.77 ± 0.29

anayake. One is that there was some problem with the analytical procedure used in the original report. Aluminium measurements are difficult to make and errors are extremely common³. Second, Tennakone and Wickramanayake might have used a different form of fluoride such as hydrogen fluoride. Third, there might be some differences in the aluminium pans. However, we chose many different types and styles of cookware and detected no significant differences in the degree of aluminium leaching.

In view of the concern that the report by Tennakone and Wickramanayake has raised among those interested in water fluoridation, we would like to emphasize that we cannot reproduce the results in that report.

J. SAVORY
J.R. NICHOLSON
M.R. WILLS

University of Virginia Medical Center,
Departments of Pathology, Biochemistry
and Internal Medicine,
Charlottesville,
Virginia 22908, USA

1. Tennakone, K. & Wickramanayake, S. *Nature* **325**, 202 (1987).
2. Brown, S., Berthoff, R.L., Wills, M.R. & Savory, J. *Clin. Chem.* **30**, 1216 (1984).
3. Cornelis, R. & Schutysse, P. *Contr. Nephrol.* **38**, 1 (1984).

Cosmic strings and celestial objects

SIR—In a recent paper¹, Tassie has suggested that celestial objects have evolved through a hierarchical breaking of rotating pieces of cosmic string and that this would account for the angular momentum (J) for a wide variety of astronomical objects being proportional to the square of their masses, (M), the constant of proportionality moreover being comparable with that of some string theories of particle physics. So one has a relation $J = KM^2$, K being a universal constant whose value from an analysis of astronomical objects of different classes is $\sim 10^{-15} \text{ g}^{-1} \text{ cm}^2 \text{ s}^{-1}$ and similarly for strings $J_s \approx K_s M^2$ with K_s being related to the string tension μ through $K_s = c(2\pi\mu)^{-1}$. For agreement with the astronomical value, that is $K = K_s$, this would imply a string tension of $\mu \approx 6 \times 10^{24} \text{ g cm}^{-1}$. It must first be pointed out that it is not necessary to invoke exotic configurations such as strings to understand the apparent $J = KM^2$ relation for celestial objects. In fact it can be shown² that just such a relation with a dimensional constant K close to the observed value, would arise naturally from considerations of the basic physics involved in the structure of various astronomical objects such as galaxies and stars. The relation in fact is an upper envelope relating maximum J to M^2 . To within a factor of order unity, systems which are gravitationally bound have typical velocities $(V/c) \approx (R_G/R)^{1/2}$, R_G being the gravitational radius

$(2GM/c^2)$. If $V^2/R \approx GM/R^2$ and $J = MVR$ we can write $J \approx (G/V)M^2$ which for $V \approx \text{constant}$ would imply an apparent $J \sim M^2$ relation. It is to be noted that the range of V or equivalently $(R_G/R)^{1/2}$ is rather narrow for astronomical objects, (as compared to the range of M), with $10^{-3} c$ being most typical. (Velocity V well below $10^{-4} c$ would be very difficult to observe and relativistic objects are rare.) Thus for $V \approx 10^{-3} c$, $K \approx 10^{-15}$, which is the empirical result seen.

Apart from the fact that there is no need to invoke strings to account for the $J = KM^2$ relation, the value for the string tension μ , which the observed value of K would imply is unacceptably high. The reason being that the dominant energy loss mechanism for strings is by gravitational radiation with the power loss given by $P = \beta G\mu^2 c$, $\beta \approx 10$. This rate of energy loss is independent of the size of the string (it depends only on the tension μ). For a string of radius R this implies a lifetime $t \approx RC/\beta G\mu$. For galaxy-size loops $R \approx 1 \text{ kpc}$, and with the above value of μ , this would give a lifetime of less than one per cent of a Hubble time, that is, they would not last long enough to form galaxies. A more stringent constraint³ (C.S., in preparation) obtained by noting that the presence of an extra component (of gravitational radiation) or even less than 5% of the total radiation background at the epoch of nucleosynthesis can affect element abundances, give $\mu < 10^{20} \text{ g cm}^{-1}$. This makes it improbable that the value of μ determines the constant of proportionality in the $J-M^2$ relation; that is, the value of μ implied from this relation being incompatible with string models of galaxy formation.

C. SIVARAM

Indian Institute of Astrophysics,
Bangalore 560034, India

1. Tassie, L. J. *Nature* **323**, 40-43 (1986).
2. Sivaram, C. in *Relativistic Astrophysics and Cosmology* (ed. de Sabbata, V.) 228-246 (World Scientific, Singapore, 1984).

TASSIE REPLIES—Defining a velocity V by $V = GM^2/J$, then, as Sivaram points out, V constant would imply an apparent $J = KM^2$ relation. But this raises the question of why V should be so constant for astronomical objects. The string picture¹ explains the constancy of both J/M^2 and V . Sivaram's V is a measure of the velocity in gravitationally bound systems, such as binary stars, but is not simply related to the rotational velocity of a planet such as Jupiter with a rotational velocity at the equator of $4 \times 10^{-5} c$. J/M^2 for Jupiter, however, is comparable with that of flat galaxies¹.

Sivaram's second point, concerning the effect of gravitational radiation, is important as it yields a bound on the time for string fragmentation. Witten² has pointed out that Type I superstrings fragment too

fast to contribute cosmologically significant gravitation radiation. As the theory of the superstring includes gravitational radiation, which is the emission of massless closed loops by the string, and because the thermodynamic treatment of string fragmentation should be independent by the detailed properties of the string, the conclusion¹ remains that the rotating string is more likely to fragment into massive pieces than any other change in its motion, including the emission of gravitational radiation. So for both Type I and Type II superstrings, the rotating string will fragment before it can emit an appreciable amount of gravitational radiation.

Sivaram's comment that superstrings would not last long enough to form galaxies would apply if superstrings formed galaxies by condensing matter around them as envisaged for galaxy formation by the cosmic strings of Kibble, but is irrelevant to the formation of galaxies by fragmentation of the superstring, where the entire mass of the galaxy comes from the mass of the superstring. As there is no ordinary matter until the hierarchical fragmentation of the superstring is completed, ordinary astrophysical processes, including nucleosynthesis, do not occur until after fragmentation.

Sivaram overestimates the life of a galaxy-forming superstring, as the string has the mass of the galaxy but a much smaller size. A superstring with the mass of a galaxy is about 1 pc in length and so has a lifetime less than 3,000 years.

Superstrings cannot form galaxies by accretion of ordinary matter, but can form galaxies by fragmentation.

L. J. TASSIE

Department of Theoretical Physics,
Research School of Physical Sciences,
The Australian National University,
GPO Box 4, Canberra,
Australia

1. Tassie, L. J. *Nature* **323**, 40-42 (1986).
2. Witten, B. *Phys. Lett.* **B153**, 243-246 (1985).

Run don't walk

SIR—The conclusion of Ker *et al.* (*Nature* **325**, 147-149; 1987), that the arch of the human foot stores enough energy to make running more energy efficient, may have been anticipated by Sir Charles Bell. In his book *The Hand its Mechanisms and Vital Endowments as Evincing Design*, published in 1833, he wrote that "The spring of the foot and the toe . . . gives elasticity and rapidity in running, dancing and leaping". I note that Ker *et al.* have, as yet, only studied running.

J.D. COLE

Wessex Neurological Centre,
Southampton General Hospital,
Shirley, Southampton SO9 4XY,
UK

High stakes at the frontier

D. H. Perkins

Nobel Dreams: Power, Deceit and the Ultimate Experiment. By Gary Taubes.
Random House: 1987. Pp. 261. \$19.95.

IT WAS inevitable that, sometime, a book would appear seeking to tell the inside story of the discovery of the charged W^+ and W^- and neutral Z^0 particles; these are the carriers of the so-called weak force which, together with electromagnetism, the strong nuclear force and gravity constitute the known fundamental interactions in nature. Their discovery provided crucial proof of the validity of a theory unifying the weak force with electromagnetism, and gained the 1984 Nobel Prize for two physicists, Carlo Rubbia and Simon van der Meer, at the international accelerator laboratory, CERN, near Geneva. The W and Z particles were created in head-on collisions of energetic protons with antiprotons circulating in opposite directions in a giant magnetic storage ring, 6 km in circumference, located underground at CERN. Van der Meer's contribution was to accomplish the daunting task of making, storing and accelerating enough antiprotons that the W and Z particles, created in only 1 per billion collisions, might be detected in sufficient numbers. Rubbia's role was to conceive and push through the entire project, and to demonstrate the existence of the particles using a sophisticated 2,000-ton detector called UA1.

The book is almost entirely centred on Carlo Rubbia, a flamboyant and brilliant physicist, a man who drove himself and his co-workers mercilessly. It is based on Taubes's sojourn of some months in CERN and on numerous interviews with physicists in the field, both from Europe and the United States. It is a racy, revealing and very readable account of opportunism in big science, and of the frequently violent interactions between personalities and factions involved in a 150-strong international collaboration of the best brains in the business — people who were playing for high stakes, for the sort of scientific breakthrough that might come once in a lifetime.

High energy physics had been dominated by the United States since the end of the Second World War, at least until the mid-1970s. In Europe, the focus of such research was CERN, where the accelerators ran with the precision of Swiss watches. But time and again, the leaner and fitter American laboratories beat CERN to the draw. The discovery of the W and Z in 1983 marked a high point in the gradual shift of the balance towards Western Europe. The shock effect in the

American science community was exemplified by the *New York Times* headline "Europe 3, US not even Z zero". It was a scientific triumph — and a desperately needed one — for CERN.

There were, in fact, two competing experiments running simultaneously at the CERN collider: UA1 and the more modest and somewhat less-well-instrumented UA2, which got comparable

CERN

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Carlo Rubbia — flamboyant and brilliant.

numbers of W and Z events but receives scant mention in Taubes's book. UA1 was a mammoth and ingenious experiment operated by gifted people, but it was only one experiment, and it could have been wrong. After all, Rubbia himself had previously taken part in more than one wrong experiment, and these mistakes had been highly visible and well publicized, because that was his style of operating. It is worth recording that the other major scientific success at CERN, the discovery of neutral currents in 1973 — and an essential step on the path to the collider project — had not been universally accepted at the time (and least of all in the higher echelons of CERN) because it had not been independently confirmed. Worse, a rival neutrino experiment at Fermilab, Chicago, had initially been in conflict with it (but turned out to have been wrong). Because Carlo Rubbia himself, in his role of professor at Harvard, had been a member of that ill-fated venture, he of all people would be the first to acknowledge the great debt he owed to UA2 for providing the much

needed confirmation of the discovery of W and Z .

What comes through clearly in Taubes's book is that the scale and expense, and the consequences of success or failure in big science projects, place great pressures on the scientists at the sharp end to deliver the right goods at the right time, and are not always calculated to bring out the best characteristics of the human species. This is at odds with the one-time popular image of basic research carried out by tiny bands of dedicated people in back rooms, quietly beaver away in their selfless search for the ultimate truths of nature. Nevertheless, all of the people on the UA1 and UA2 experiments were there because they were driven by the physics. Personal ambitions, power and glory, politics and backbiting also came into it, and, of course, these things reveal the normal human strengths and weaknesses, make the headlines and sell popular accounts; but when the dust settles, they are forgotten and the solid scientific achievements remain.

Taubes also traces some of the developments in the CERN UA experiments since the W and Z were first observed, in particularly favourable modes of decay, into electrons and muons and neutrinos. They have not been found in more prolific decay modes, which we know must be there, because other physical processes obscure the signal. But that very "background" has turned out as a bonus in another way, giving important results — not originally foreseen — on the nature of the forces between quarks in the proton. This is not even mentioned in the book. Other new phenomena had also been predicted by theory — the existence of the top quark and of supersymmetric particles. Here the experimenters have not so far been successful, despite enormous efforts. Initial evidence for the existence of a new quark with 40 GeV mass has not been confirmed, and early indications for supersymmetric particles (the so-called monojets) have melted into the background. That is what research at the frontiers is all about. It is a cautionary tale. You try hard but you just can't win them all.

For all the human interest, the real message of Taubes's book is that it describes a triumph of modern technology and scientific ingenuity against enormous odds, one which enlarged and enriched our view of the material world. Not least, that and the Nobel award which followed were a tangible demonstration that a dozen European states could sink their differences and work together in a huge and complex enterprise towards a common scientific goal. □

D. H. Perkins is a Professor in the Department of Nuclear Physics, University of Oxford, Keble Road, Oxford OX1 3RH, UK.

The cosmological racket

John D. Barrow

Inflationary Cosmology. Edited by L. F. Abbott and So-Young Pi. *World Scientific*:1986. Pp.697. Hbk £77.55, \$89.70; pbk £31.80, \$36.80.

DURING the past few years there has grown up an unwelcome addition to the scientific publishing business. Armed with a photocopier and a reasonably stocked periodicals' library, anyone can now create a large book in an afternoon.

Here's how you play. First, run through the literature of your subject picking out 50 or 60 interesting papers (if you know this literature reasonably well already you can skip this stage and simply turn out your files). Next, take two photocopies (one for the publisher and one for you to keep — you don't want to end up having to buy the "book" do you?). Now shuffle them into a set of seven or eight related subjects. Sit down and write a page or two of introduction to each and then call it a "chapter". Finish in time to catch the afternoon post to the publishers; they do the rest. The result in this case is a book of offprints which you could have photocopied for yourself, or obtained free of charge from the contributing authors, but which sells in hardback for no less than seventy seven pounds in Britain or a shade under ninety dollars in the United States.

What can one say about books such as this? Only the copyright holders of the articles have it within their power to prevent their appearance. Even if inexpensive I would still argue that they are a bad thing. They encourage the beginning student of the subject at whom they are aimed to avoid contact with the research literature. Thereby the student fails to find interesting papers by chance and fails to learn how to use the scientific literature systematically and serendipitously. The editors of the collection seek out only papers that support a particular approach or theme and create the artificial impression that a very speculative and exciting research area is literally a closed book.

Of course, some of these defects can be remedied by good contextual analyses and connecting discussions by the editors. Here, little effort has been put into this potentially mitigating feature. The introductions to the chapters are only about two sides in length on average, and are too brief to describe adequately the complexities involved — many of the papers were

written at quite different times and for different reasons, and even come to conflicting conclusions. The few pages of introduction that do appear are not always satisfactory, especially in the early parts of the book (for example, the Friedman equation is given incorrectly in the opening pages). Confusion is also caused by the continual assertion that the theory of an inflationary universe predicts that today the Universe should have density *equal* to the critical density. Whereas, in fact it predicts only that the density must now be very close to the critical density. This distinction is crucial. If the density were equal to the critical value today it would always have been, and there would have been no role for inflation to play in explaining the fact.

A more serious misjudgement is in the selection of observational evidence in the chapter on dark matter. Unfortunately, the sole observational paper discussing the evidence for dark matter in spiral galaxies is devoted to optical rotation curves, even though it is now well known that these observations do not require dark matter for their explanation. The key observational data are the 21-centimetre observations of spiral galaxies per-

formed by radio astronomers. It is this work that establishes the existence of dark matter well beyond the optical extent of the Galaxy and is the cornerstone of all observational evidence for dark matter. Not one report of these data is included.

The good points of the collection are of course the papers themselves, some worthy inclusions having been written by the editors and their collaborators. The combination of authors from observational astronomy, particle physics and astrophysics illustrates the healthy and ever-expanding state of modern cosmology. But anyone working seriously in the subject will already have read 90 per cent of these papers and will be familiar with all of them.

Librarians whose budget cannot keep pace with the other inflationary universe should be aware that their periodicals' collection already contains the contents of this book. Individuals working in Third World or soft currency countries can obtain them free by sending reprint requests to the authors of the individual papers. □

John D. Barrow is a Lecturer at the Astronomy Centre, University of Sussex, Brighton BN1 9QH, UK.

Peculiar progress

Tony Davies

Paradoxes in Immunology. Edited by Geoffrey W. Hoffman, Julia G. Levy and Gerald T. Nepom. *CRC Press*:1986. Pp.336. \$114.50 (United States), \$131.50 (elsewhere).

It is a common problem among teachers that ignorance of a particular phenomenon is difficult to impart; the rapt student often wishes only for the comfortable wash of supposedly hard facts. Research workers, on the other hand, must start from the premise that the contemporary array of knowledge is inadequate and that it is their duty to add to or change it. Researchers teach occasionally, and one serious consequence is that, perhaps affected by their teaching activities, they often invent an authoritative dogma which can impose constraints on the process of discovery. These constraints are often referred to as working hypotheses, and the intent of their inventors sometimes seems to prove them come what may.

In biology, because the working hypotheses are commonly at variance with the truth, paradoxes arise. Such paradoxes in immunology are here collected together in the form of a book. It is mildly surprising that they have to be viewed as irrationalities rather than small flaws in the thin veneer of our knowledge.

The three editors wrote to 150 immu-

nologists asking them to contribute their own pet anomaly. Forty said they would. Twenty-seven actually did. The papers were in the main assembled in the same format; first, exposition of the paradox, then a statement of the two apparently contrasting sets of facts and a brief discussion to follow. The final versions were then sent out to various high priests in immunology, nominated by the authors, for their comments. In some instances the original authors have exercised a brief right of reply.

It all sounds the very stuff of science — proposition, argument and counter-argument. What could be better? As presented it is, I regret to say, not easy to understand. I doubt whether the editors could explain more than a small fraction of the material in the book to their own students. The fault is not entirely theirs in that contemporary immunology is a morass of partly digested and ill-coordinated facts, many of which are almost surrealistic in the indirectness of their impact.

For all this, the intention of the editors was a worthy one. Their aim, they write, was to enhance the rate of progress in immunology. They see a collection of paradoxes coming out on a regular basis and encourage the submission of contributions. The idea ought to work in time, but in practice much better editing, more independent refereeing and more plain speaking will be required. Biologists invent their own jargon which, not intentionally, has the effect of isolating particu-

• A second edition of *Ideals and Realities: Selected Essays of Abdus Salam* has been published by World Scientific. Price is hbk £37.75, \$47.20; pbk £23.85, \$29.80. The first edition was reviewed in *Nature* 312, 216 (1984).

lar branches of the tree of learning. If highly specialized books such as this are to be useful, a more serious attempt has to be made to cut through the jargon and establish stronger lines of communication.

The content of the book ranges from the smell of histocompatibility antigens (this is a joke not a paradox, if all that has been written about histocompatibility antigens relates to an outbreeding system rather than say associative recognition at the cellular level), through the lack of class I type histocompatibility antigen polymorphisms in Syrian hamsters (how can they survive when such polymorphisms are deemed to be almost a main staff of life), to the enhancement of skin allografts by prior blood transfusion (it should immunize). The astonishing fact that many tumours can remain dormant for years is seen as a contradiction of the notion that the immune response is designed in part to see tumours off. I am inclined to agree with the commentator (Richmond Prehn) who observes laconically "the fact of dormancy constitutes a strong argument against the thesis that the immune response is the major homeostatic mechanism vis à vis cancer". In other words, the apparent paradox strikes at the root of the working hypothesis, which clear indication is, in turn (and paradoxically) largely ignored by the authors of the chapter in question.

Many of the chapters are concerned with the paradigm of the immune response as a defence mechanism, which indeed it may be. But interested immunologists should pay attention to at least the title of a paper quoted on p.114 ("Thomas, L. 'Symbiosis as an immunologic problem' in *The Immune System and Infectious Diseases*, Milgrom, F. and Neter, E., eds, S. Karger, Basel, 1975, 2"). If in fact the immune response is not by design rejectional but accommodatory, then at least a proportion of the paradoxes vanish overnight — but that's my prejudice!

The book merits careful reading. It illustrates the ease with which immunologists remain encapsulated in the intellectual straitjackets that they themselves have put on, and the almost catechismal rigidity with which they regard working hypotheses. If the editors were to publish one of these paradoxes a month in some more popular and less-expensive conveyance of scientific news, if they were to insist both in advance and by editing that the exact nature of the apparent paradoxes was intelligibly displayed, if they were to read again the fable of the Emperor's clothes, then they just could produce something of value in these uneasy days of grant-supported science. □

Tony Davies is Professor of Immunobiology and Deputy Director at the Institute of Cancer Research, Haddow Laboratories, Clifton Avenue, Belmont Sutton, Surrey SM2 5PX, UK.

Empirical insight

Peter E. Hodgson

The Neglect of Experiment. By Allan Franklin. Cambridge University Press: 1987. Pp.290. £30, \$42.50.

Few philosophers of science have spent their formative years struggling with recalcitrant apparatus and trying to decide whether some observations really are significant or should be attributed to statistical fluctuations or equipment malfunction. Most of them started as theoretical physicists or mathematicians before they turned to philosophy, and so tend to see the history of science as the battleground of competing theories, with only occasional reference to experiments. Such an approach ignores a whole series of questions bearing on the understanding of science: Why do scientists decide to do a particular experiment? How do they know when the results are significant? Why are some anomalies ignored and others hailed as decisive discoveries?

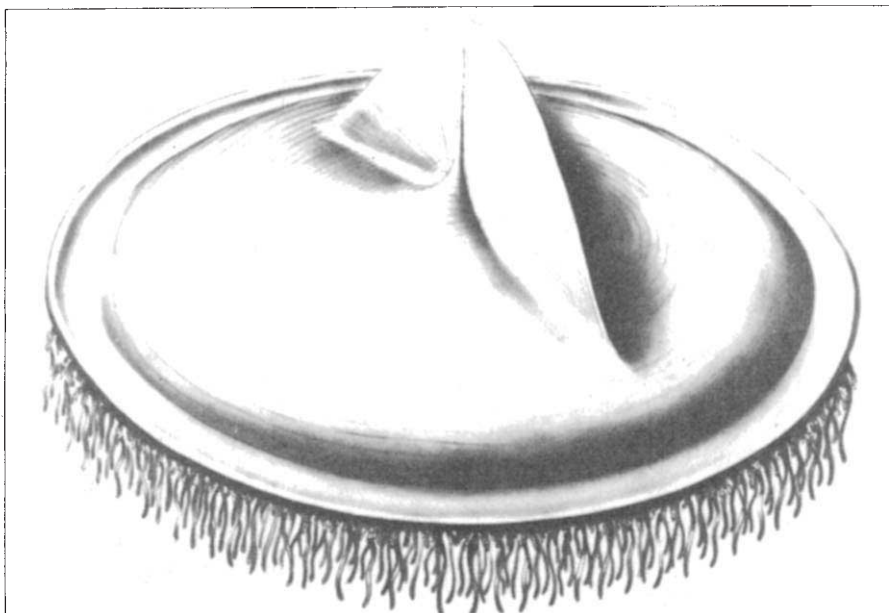
Professor Franklin began as an experimental high-energy physicist and so has practical experience of these matters. He now works on the history and philosophy of science, and so is well-equipped to redress the balance. Problems about the role of experiment have been discussed by philosophers, but usually in the abstract. Franklin adopts the empirical approach, analysing in detail several decisive experiments in elementary particle physics. He has uncovered a wealth of fascinating detail that supports the validity of the methods used and shows the inadequacy of many philosophical accounts. As far as

possible, he has used the original data, the laboratory notebooks and the personal recollections of the scientists who did the work concerned.

His first example of a crucial experiment is that leading to the discovery of parity non-conservation. The experimental evidence was overwhelming; it solved the τ - θ puzzle and is of vital importance in elementary particle and nuclear physics. Subsequently it was confirmed by a wide range of different experiments. Oddly enough, convincing results of parity-violating experiments were published in 1928 and 1930, but the implications were not recognized. Franklin suggests that this was partly because of the lack of an appropriate theoretical context at the time, and partly because the issues at stake were overshadowed by other contemporary arguments.

His next example is the experiment of Cronin, Fitch and colleagues that refuted the concept of the invariance of physical processes under combined space inversion and particle-antiparticle interchange. This was a complicated experiment that required exhaustive tests of all conceivable alternative hypotheses. Not all of these are reported in the original papers, which therefore tend to give the unwary reader an inadequate impression of the validity of the procedure.

A simpler example is provided by Millikan's oil-drop experiment that established the quantization of charge and gave an accurate value for the charge of the electron. Franklin examined Millikan's notebooks and repeated his calculations. He found that although Millikan said that he had published all his results, he had in fact been somewhat selective. One series of measurements did not give the "expected"



Sailing back — reconstruction of the palaeozoic hydrozoan Plectodiscus discoideus. The picture is taken from Problematic Fossil Taxa, edited by Antoni Hoffman and Matthew H. Nitecki and published by Oxford University Press. Price is £60, \$69.

result and was rejected: "Not an oil drop" is written in the notebook. Other drops were rejected for a variety of reasons. But although Millikan indulged in some trimming, the reanalysis confirms the validity of his procedure and the accuracy of his results.

These case histories, and a few others more briefly described, provide valuable material for philosophical reflection. They strongly support the scientist's belief that his experiments tell us about the world, and that crucial experiments can be carried out. This is worth saying, particularly as some philosophies of science throw doubt on it. Thus the kuhnian account of scientific revolutions stresses that the incommensurability of the terms in competing paradigms prevents experimental comparison. Franklin, however, shows that although experiments are theory-laden they can be described in theory-neutral terms and can provide a crucial test of competing paradigms. This is illustrated by an account of such an experiment for newtonian and einsteinian mechanics, an exemplary case used by Kuhn. An historical example is Lummer and Pringsheim's experiment on the spectrum of black-body radiation that led to Planck's quantum theory.

Franklin also discusses the Duhem-Quine problem, namely that it seems that a theory can never be definitely refuted

because it is always possible to modify some auxiliary assumptions to bring it into line with experiment. How do we know, for example, that in the Wu parity non-conserving experiment something did not invert somewhere else to maintain parity conservation? There is no answer to this type of scepticism, but it is irrelevant to science. Scientists have developed their own ways of dealing with the problem of auxiliary assumptions, and Franklin analyses them in terms of bayesian confirmation theory. Essentially, this says that "we have degrees of belief in statements or hypotheses and that these degrees of belief obey the probability calculus". Franklin shows that they make good sense of his chosen examples of crucial experiments.

Franklin's analyses show clearly the overall reliability of experimental science, confirming and justifying the methods instinctively adopted by working scientists. Scientists seldom study the philosophy of science before embarking on their researches, and arguably this is just as well. But certainly a deeper insight into what they are doing is valuable, and this has now been provided. □

Peter E. Hodgson (who began his career in physics as an experimentalist) is Head of the Nuclear Physics Theoretical Group, Nuclear Physics Laboratory, University of Oxford, Keble Road, Oxford OX1 3RH, UK.

A rare discipline

James D. Thomson

Conservation Biology: The Science of Scarcity and Diversity. Edited by Michael E. Soulé. *Sinauer/Blackwell Scientific*: 1986. Pp.584. Hbk \$50, £40; pbk \$27.95, £23.

Books that attend the emergence of new fields have double agendas. At one level, these 25 chapters by 45 authors provide a broad review and partial synthesis of conservation biology. But 'conservation biology', which involves the self-conscious transfer of principles from the pure realm of ecology and evolution to the applied realm of preservation and repair, is a juvenile and polymorphic science. This book will join its predecessor (*Conservation Biology: An Evolutionary - Ecological Perspective*, edited by M. E. Soulé and B. A. Wilcox, and published by Sinauer/Blackwell Scientific in 1980) as one of the founding documents of the field. Unlike treatises in mature fields, therefore, it shares the important responsibility of defining a discipline and determining its future character. Fortunately, it is up to the task.

That future character will be strongly influenced by resolution of the tension

between academic scientists and preserve managers. To oversimplify greatly, one important barrier is the unwillingness or inability of academics to devote themselves to applied problems. The work is tough, the rewards unconventional. A second is the unreceptive attitude of many managers, who often find what advice they do get from academics to be unspecific, unrealistic, equivocal and arrogant. Although Soulé has chosen authors mostly from the academic side, he has strengthened the book immeasurably by addressing the pure-applied tension head on. Starting with Chapter 1, "Conservation Biology and the 'Real World'", and continuing through his overviews of each of the six sections, Soulé serves as the conscience of the book as he repeatedly points out implications for management, highlights those questions most likely to seduce academics into entering the fray, and tells both sides how to behave.

Soulé's overviews further improve the book by linking and clarifying the individual chapters; it is extremely helpful, for example, to have an explanation of the distinction between Ledig's and Templeton's uses of the term 'genetic coadaptation', or of the functional correspondence of the different taxonomies of rarity used by Cody, Rabinowitz *et al.*, and Hubbell and Foster.

Most of the chapters are meaty and

provocative. Unsurprisingly, they span a great range from cautious optimism to pessimism, concerning both the salvation of natural diversity and the success of conservation science. Consistent with the generally diminished expectations for a concise theoretical ecology in the face of natural complexity and idiosyncrasy, the ecologists see less hope for broadly applicable management principles than do the geneticists. For example, while extracting some potentially useful generalizations about community stability, the relatively optimistic theoretician Pimm needs to hedge on several fundamental points — for example, a successful species introduction could cause "no, few, or many species losses" (p. 325). Speaking as a field ecologist and a preserve manager who really has to face edge effects, Janzen offers few general prescriptions other than the apt, but rather bleak, "Know the natural history of your organisms well enough that you can both anticipate their interactions and know at what distance their populations can be perturbed by a preserve edge" (p. 303). On the other hand, Templeton, a population geneticist, shows that laboratory experiments with *Drosophila* can provide the insights needed to set up breeding management schemes for genetically vulnerable lizards and gazelles. Biology appears to become more predictable as one gets closer to the DNA (but cleansing a population of bad gene combinations is a much tidier problem than guaranteeing it a place to live).

Among the chapters on particular habitats, the treatments of tropical deforestation (Myers) and African desertification (Le Houéron and Gillet) are especially good — and especially depressing, because both disasters continue to be driven by human overpopulation which seems virtually unstoppable. Soulé argues convincingly, however, that scientific input of the right sort *can* ameliorate the damage to developing countries, and Diamond proves the point that even single individuals can make a difference.

Overall, Soulé's book not only explains conservation biology, it inspires one to do it. This brings us to the conservation of conservation biologists. With the unseemly rush towards molecular biology, the abandonment of traditional training grounds (such as botany departments) and the pervasive academic disdain for applied studies (except those that make money), will there be enough habitats in tomorrow's universities to support a viable population of conservation scientists? As Soulé puts it (p. 308), "The big question is, who will take the lead in developing a management curriculum for the 21st century?" □

James D. Thomson is an Associate Professor of Ecology and Evolution at the State University of New York, Stony Brook, New York 11794, USA.

Combined La–Ce and Sm–Nd isotope systematics in petrogenetic studies

T. Tanaka*, H. Shimizu†, Y. Kawata† & A. Masuda†

* Geological Survey of Japan, Higashi 1-1-3, Yatabe, Ibaraki 305, Japan

† Laboratory for REE Microanalysis, Department of Chemistry, University of Tokyo, Hongo, Tokyo 113, Japan

A theoretical model is presented for the coupled evolution of cerium and neodymium isotopes in terrestrial rocks. Data presented for samples ranging from mid-ocean-ridge basalts to continental metamorphic rocks show that, while mantle-derived rocks are dispersed on a Ce–Nd diagram, continental crustal rocks form a clearly defined ‘mantle array’. As analytical techniques improve, cerium isotopes will provide valuable constraints in studies of petrogenesis and crustal evolution.

THE isotope ^{138}La decays to ^{138}Ce and ^{138}Ba by β^- -emission and electron capture, respectively, with a total half life of about 1×10^{11} yr. Tanaka and Masuda¹ successfully used the ^{138}La – ^{138}Ce decay scheme as a chronometer for a gabbro from the Bushveld Complex in South Africa. The La–Ce isotope system, coupled with the Sm–Nd system, also has potential as an isotope tracer^{1–3}. The purposes of the present study are to investigate the isotopic composition of Ce in selected terrestrial rocks and to determine the potential applications of Ce isotopes in petrogenesis and the evolution of the Earth’s crust and mantle.

Samples and analysis

Samples analysed in this study were collected from various terrestrial geological units. Basalts include mid-ocean-ridge basalts (MORB), two from the Mid-Atlantic Ridge and one from the East Pacific Rise, ocean island basalts (OIB) from Kilauea, Mauna Ulu, Iceland and Gough, island arc basalts (IAB), from Izu Oshima, Fuji and Oki, and continental plateau basalts (BCR-1, Decan). Gneisses and metagabbro which constitute the continental crust (from Africa and North America) and granodiorite (JG-1) and rhyolite (JR-1) from the Japanese Islands were also analysed.

Radiometric ages of these rocks have been obtained using various methods. We chose the oldest age, with the exception of model ages, as the assumed emplacement age for each rock sample. The ages are listed in Table 1 and are used in the following discussions. Samples without chronological data in Table 1 are younger than 100 Myr.

The experimental procedures used are almost identical to those of Tanaka and Masuda¹. About 0.3 to 1 g of each sample was decomposed. Rare-earth elements (REEs) were separated from the major elements by ion exchange using an AG50W-X8 resin column, eluting with HCl. Cerium and neodymium were then separated using an AG50W-X8 resin column with α -hydroxy-isobutyric acid (α -HIBA). The α -HIBA process was repeated for some samples with higher Nd/Ce ratio. The isotopic composition of cerium and neodymium was measured on a semiautomatic Micromass 30–54 R mass spectrometer, and abundances of lanthanum, cerium, neodymium and samarium were determined on a JEOL JMS-05RB and a Micromass 30–54 R. The tailing of peak 156 ($^{140}\text{Ce}^{16}\text{O}$) on peak 154 ($^{138}\text{Ce}^{16}\text{O}$) (~60 p.p.m. of $^{138}\text{Ce}^{16}\text{O}$) was monitored and corrected. Larger sample loads (5–10 μg), than used previously¹, cause greater tailing. Cerium isotopes were measured as CeO^+ using Re triple filaments and the observed $^{138}\text{Ce}/^{142}\text{Ce}$ ratios were normalized to $^{136}\text{Ce}/^{142}\text{Ce} = 0.01720$. The neodymium ion was measured as NdO^+ using an Re single filament and the measured isotope ratios were normalized against $^{146}\text{Nd}/^{144}\text{Nd} = 0.7219$. Results for $^{138}\text{Ce}/^{142}\text{Ce}$ and $^{143}\text{Nd}/^{144}\text{Nd}$ ratios as well as $^{138}\text{La}/^{142}\text{Ce}$ and $^{147}\text{Sm}/^{144}\text{Nd}$ ratios are shown in Table 1.

The Ce and Nd isotope ratios were normalized to chondritic values and presented using the ϵ notation as described by

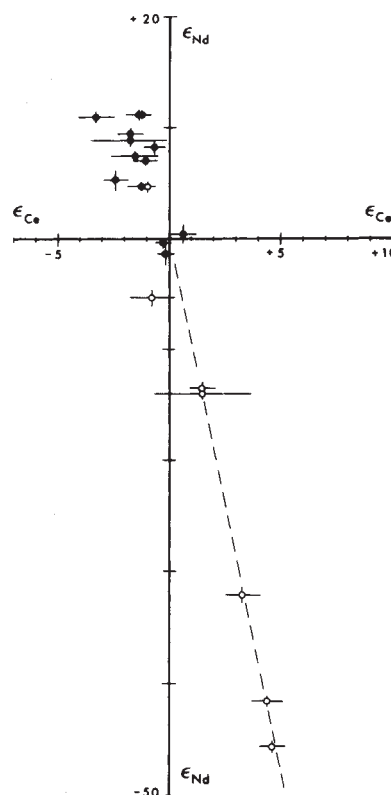


Fig. 1 $\epsilon_{\text{Ce}}-\epsilon_{\text{Nd}}$ plot of various terrestrial rocks. Open circles indicate samples from continental crust and granodiorite and rhyolite from Japanese Islands. Solid circles indicate basic volcanics. Dashed line indicates $\epsilon_{\text{Ce}} = -0.112 \epsilon_{\text{Nd}}$ (see the text).

DePaolo and Wasserburg⁷. $\epsilon(T) = (R_m(T)/R_{\text{CHUR}}(T) - 1) \times 10^4$ where $R_m(T)$ and $R_{\text{CHUR}}(T)$ refer to the isotope ratios $^{138}\text{Ce}/^{142}\text{Ce}$ or $^{143}\text{Nd}/^{144}\text{Nd}$ for a sample (m) and chondritic uniform reservoir (CHUR), respectively, at time T ($T=0$ for present). As our analysed value for the LaJolla Nd standard (0.51183 ± 2) is lower by three in the last digit than the value of LaJolla⁸, we added three to the last digit of our data in Table 1, and then, ϵ values in Table 1 were calculated. The age (T) under consideration here is the present, and the values $R_{\text{CHUR}}^{\text{Ce}}$ and $R_{\text{CHUR}}^{\text{Nd}}$ used for normalization are 0.0228527 (ref. 9) and 0.51264 (ref. 6), respectively. The $\epsilon_{\text{Ce}}-\epsilon_{\text{Nd}}$ plot is shown in Fig. 1.

Ce–Nd isotope model

To define some general features of Ce–Nd isotope systematics as shown in Fig. 1, we performed model calculations before discussing individual cases. Covariations of $^{138}\text{Ce}/^{142}\text{Ce}$ and

Table 1 La-Ce and Sm-Nd isotope data for geological samples examined

Sample	La (p.p.m.)	Ce (p.p.m.)	$^{138}\text{La}/^{142}\text{Ce}$	$^{138}\text{Ce}/^{142}\text{Ce} \pm (\epsilon)$	Sm (p.p.m.)	Nd (p.p.m.)	$^{147}\text{Sm}/^{144}\text{Nd}$	$^{143}\text{Nd}/^{144}\text{Nd} \pm (\epsilon)^\dagger$
Mid-ocean ridge basalts								
1	2.90	9.29	0.00253	$0.0228450 \pm 19(-3.4 \pm 0.8)$	3.09	8.67	0.2156	$0.51317 \pm 2(+10.9)$
2	4.12	13.36	0.00250	$0.0228494 \pm 13(-1.4 \pm 0.6)$	4.30	12.59	0.2066	$0.51318 \pm 2(+11.1)$
3	18.72	59.5	0.00255	$0.0228498 \pm 9(-1.3 \pm 0.4)$	15.15	48.4	0.1893	$0.51318 \pm 2(+11.1)$
Ocean island basalts								
4	19.44	50.4	0.00313	$0.0228503 \pm 14(-1.1 \pm 0.6)$	8.04	32.4	0.1501	$0.51297 \pm 2(+7.0)$
5	10.91	27.7	0.00319	$0.0228471 \pm 13(-2.5 \pm 0.6)$	4.86	18.64	0.1577	$0.51288 \pm 5(+5.3)$
6	2.97	7.90	0.00305	$0.0228486 \pm 38(-1.8 \pm 1.7)$	2.17	6.39	0.2054	$0.51306 \pm 2(+8.8)$
7	7.56	20.3	0.00302	$0.0228512 \pm 12(-0.7 \pm 0.6)$	3.82	13.70	0.1687	$0.51303 \pm 3(+8.2)$
8	35.3	73.1	0.00392	$0.0228522 \pm 10(-0.2 \pm 0.4)$	6.07	32.2	0.1140	$0.51254 \pm 5(-1.4)$
Island arc basalts								
9	2.25	6.65	0.00274	$0.0228485 \pm 14(-1.8 \pm 0.6)$	2.24	6.22	0.2178	$0.51309 \pm 3(+9.4)$
10	8.27	21.1	0.00316	$0.0228491 \pm 26(-1.6 \pm 1.1)$	4.14	15.50	0.1616	$0.51299 \pm 2(+7.4)$
11	29.5	65.8	0.00364	$0.0228540 \pm 12(+0.6 \pm 0.6)$	6.93	34.1	0.1229	$0.51263 \pm 5(+0.4)$
Plateau basalts								
12	9.30	24.3	0.00310	$0.0228498 \pm 12(-1.3 \pm 0.6)$	5.19	18.18	0.1727	$0.51285 \pm 2(+4.7)$
13	25.1	53.6	0.00380	$0.0228520 \pm 10(-0.3 \pm 0.4)$	6.70	28.9	0.1402	$0.51259 \pm 2(-0.4)$
Island arc acidic rocks								
14	17.84	46.3	0.00312	$0.0228505 \pm 10(-1.0 \pm 0.4)$	5.75	23.1	0.1506	$0.51285 \pm 2(+4.7)$
15	18.54	39.8	0.00378	$0.0228508 \pm 26(-0.8 \pm 1.1)$	3.61	16.6	0.1315	$0.51234 \pm 3(-5.3)$
Continental rocks								
16	38.5	78.1	0.00400	$0.0228603 \pm 20(+3.3 \pm 0.9)$	5.44	33.8	0.0974	$0.51097 \pm 4(-32.0)$
17				$0.0228561 \pm 13(+1.5 \pm 0.6)$				$0.51192 \pm 3(-13.5)$
17*	1,743	4.23	0.00334	$0.022856 \pm 5(+1.5 \pm 2.2)$	0.689	2.53	0.1647	$0.51189 \pm 4(-14.0)$
18	62.0	120.0	0.00419	$0.0228628 \pm 15(+4.4 \pm 0.7)$	8.00	44.1	0.1097	$0.51048 \pm 2(-41.5)$
19	56.5	99.8	0.00459	$0.0228633 \pm 13(+4.6 \pm 0.6)$	5.50	33.3	0.0999	$0.51027 \pm 2(-45.6)$
Meteorite			0.00306	0.0228527			0.1966	0.51264

$^{138}\text{La}/^{142}\text{Ce} = (\text{La}/\text{Ce})_{\text{weight}} \times 0.00811$ (ref. 1), $^{147}\text{Sm}/^{144}\text{Nd} = (\text{Sm}/\text{Nd})_{\text{weight}} \times 0.6049$ (ref. 6).

(1) Mid-Atlantic Ridge 22°59.14' N/43°30.90' W, Leg 46. Hole 396B 1-7-3-1B.

(2) Mid-Atlantic Ridge 23°02.63' N/45°01.05' W, dredged sample AII-92-29-1.

(3) East Pacific Rise 12°52' S/110°57' W, dredged sample E-89(D-3).

(4) Kilauea East Rift Zone, October 1977 lava.

(5) Mauna Ulu, 1973 lava.

(6) Iceland lava flow from Trolladyngja shield volcano, SA 2302.

(7) Iceland prehistoric lava flow probably from Askja caldera, SA 2407.

(8) Gough Island ankaramite ALR26G.

(9) Izu Oshima, tholeiite, 1950-1951 lava, Geological Survey of Japan standard rock JB-2.

(10) Fuji Volcano, high alumina basalt, Geological Survey of Japan standard rock JB-3.

(11) Oki Island, Japan, alkali-olivine basalt T81050307 (Saigo basalt).

(12) Decan tholeiite MAW-28.

(13) Columbia River Basalt, U.S. Geological Survey standard rock BCR-1.

(14) Wada-Toge rhyolite, Geological Survey of Japan standard rock JR-1.

(15) Sori granodiorite, Geological Survey of Japan standard rock JG-1.

(16) Broderic Falls, Kenya gneissose hb. biot. granodiorite S-83002.B (2,100 Myr, unpublished).

(17; 17') Bushveld gabbro KS-70030821 data of 17' are from ref. 1 (2,050 Myr, refs 1 and 4).

(18) Typical Morton gneiss (Rainbow Granite), Minnesota R19565 (3,000 Myr, ref. 5).

(19) Grey tonalitic gneiss, Minnesota R19569 (3,500 Myr, ref. 5).

* Raw data obtained from the Geological Survey of Japan.

† Normalized against La Jolla Nd standard 0.51186 (see text).

‡ Errors are $2\sigma_m$.

$^{143}\text{Nd}/^{144}\text{Nd}$ are computed for selected chondrite-normalized REE patterns (which are shown in Fig. 2), and different time integrations. The chondrite-normalized REE patterns in Fig. 2 are chosen rather arbitrarily. L_1 shows a slightly La-enriched logarithmically linear REE pattern. The Bushveld gabbro examined here and Amitsoq banded granodiorite gneiss 155803¹⁰ in Greenland are examples of a pattern like L_1 . L_3 shows a strongly light REE-enriched pattern, as observed for the Broderic Falls granodiorite examined here and for the Northern Light gneiss¹¹. L_2 shows a moderately enriched REE pattern as observed for a composite of North American shales¹². On the vertical scale of Fig. 2, absolute value is not an important factor. S_1 shows a strong depletion in light REEs (LREEs). Some Mid-Atlantic Ridge gabbros from the Atlantis collection A-150 correspond to this type¹³. Typical MORBs have REE patterns as represented by S_2 or S_3 .

$^{138}\text{Ce}/^{142}\text{Ce}$ and $^{143}\text{Nd}/^{144}\text{Nd}$ variations of L_1 - L_3 and S_1 - S_4 are calculated for different geological integration times. Since the variations in $^{138}\text{Ce}/^{142}\text{Ce}$ and $^{143}\text{Nd}/^{144}\text{Nd}$ are small, the data are presented normalized to CHUR as described above. It is assumed that a fractionation event took place T years ago to change the $^{138}\text{La}/^{142}\text{Ce}$ and $^{147}\text{Sm}/^{144}\text{Nd}$ ratios from the chondritic ratios to those presented in Fig. 2; $\lambda_{^{147}\text{Sm}} = 6.54 \times 10^{-12} \text{ yr}^{-1}$, $\lambda_{^{138}\text{La}} = 2.58 \times 10^{-12} \text{ yr}^{-1}$ and $\lambda_{^{143}\text{Nd}} = 4.59 \times 10^{-12} \text{ yr}^{-1}$ are used for this calculation where λ is the radioactive decay constant. Results are shown in the $\epsilon_{\text{Ce}} - \epsilon_{\text{Nd}}$ plot (Fig. 3). By definition, the values of ϵ_{Ce} and ϵ_{Nd} are zero

for CHUR, and the absolute magnitudes of ϵ values are larger for systems with strongly LREE-enriched patterns (L_3) than for the gently La-enriched REE pattern (L_1). A remarkable feature is that the trends of variation with time T for L_1 - L_3 are almost identical despite the different gradients of the REE patterns as shown in Fig. 2.

On the other hand, the $\epsilon_{\text{Ce}} - \epsilon_{\text{Nd}}$ variation trends for the LREE-depleted patterns (S_1 - S_4) diverge from each other. The strong LREE-depleted pattern (S_1) shows the largest $\epsilon_{\text{Ce}} - \epsilon_{\text{Nd}}$ variation with time integration, and has the smallest $\epsilon_{\text{Ce}}/\epsilon_{\text{Nd}}$ ratio. ϵ_{Ce} and ϵ_{Nd} for the slightly LREE-depleted pattern (S_4) do not vary so much, but the $\epsilon_{\text{Ce}}/\epsilon_{\text{Nd}}$ ratio is larger than that of the strongly LREE-depleted pattern (S_1). Samples with the same REE curvature and different time integrations are plotted on a single line (the solid line in Fig. 3) corresponding to the degree of curvature of their REE patterns. The samples with different REE curvature but with the same geological time integration define a single isochron curve (shown as dashed lines in Fig. 3). Thus, these theoretical considerations lead us to conclude that, for the LREE-enriched patterns (L), the points appear convergent along a line, while for the LREE-depleted patterns (S), the points appear divergent.

Discussion

The old crustal rocks plotted in Fig. 1 define a linear array, $\epsilon_{\text{Ce}} = -0.112 \epsilon_{\text{Nd}}$, passing through the origin of the diagram (CHUR-point). The theoretically calculated inclinations of L_1 ,

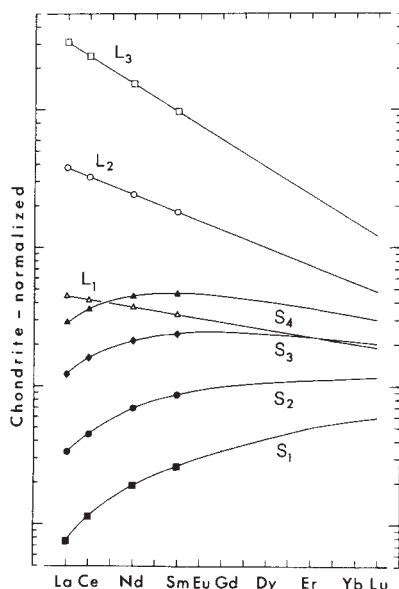


Fig. 2 Chondrite-normalized REE patterns assumed in the calculation of ϵ_{Ce} and ϵ_{Nd} variations in Fig. 3. Absolute ordinate scale is not an important factor.

L_2 and L_3 in Fig. 3 are -0.097 , -0.099 and -0.110 , respectively. The absolute values of the theoretically obtained inclinations are slightly lower than the observed value -0.112 . As shown in Fig. 3, the inclination is determined almost entirely by the ratio of the decay constants, that is, $(\lambda_{\beta}^{138}\text{La})/(\lambda_{\alpha}^{147}\text{Sm})$, and is not influenced so much by the $(\text{La/Ce})/(\text{Sm/Nd})$ ratio. This suggests that the value of the decay constant $\lambda_{\beta}^{138}\text{La} = 2.58 \times 10^{-12} \text{ yr}^{-1}$ used in this paper may be slightly lower than the actual decay constant.

The Morton gneiss has a high La/Ce ratio like L_3 , while the Bushveld gabbro has a lower La/Ce ratio like L_1 . Theoretically, these two data points are expected to be plotted on different lines. Mutual divergence of the lines, however, is quite small as shown in Fig. 3 and the current accuracy of our $^{138}\text{Ce}/^{142}\text{Ce}$ measurement is not sufficient to distinguish the difference.

The fractionation history of the REEs before the formation of igneous rocks is one of great interest in petrogenesis. Although Sm-Nd isotope systematics has provided time-integrated information on the Sm/Nd ratio by initial isotopic composition at the time of emplacement, so far we have not had such constraints on whole REE patterns. Variations in the REE pattern for the LREEs (La, Ce, Nd) predominate over those for heavy REEs. Provided that information is available for the La/Ce ratio, the present Ce isotopic composition and the formation age, one can calculate an initial Ce isotopic composition of source materials from time-integrated information on the La/Ce ratio.

We assume that the old crustal rocks under consideration were emplaced T years ago as indicated by radiometric data (Table 1). The initial isotopic composition of cerium ($^{138}\text{Ce}/^{142}\text{Ce})_T$ can be calculated as follows:

$$\left(\frac{^{138}\text{Ce}}{^{142}\text{Ce}}\right)_T = \left(\frac{^{138}\text{Ce}}{^{142}\text{Ce}}\right)_P - \frac{\lambda_{\beta}}{\lambda_{\beta} + \lambda_{\text{EC}}} [e^{(\lambda_{\beta} + \lambda_{\text{EC}})T} - 1] \left(\frac{^{138}\text{La}}{^{142}\text{Ce}}\right)_P$$

Suffixes T and P refer to the assumed emplacement age and the present, respectively. Chondritic $(^{138}\text{Ce}/^{142}\text{Ce})_T$ values can be calculated for the corresponding ages of crustal rocks under consideration. Then, $\epsilon_{\text{Ce}}(T)$ values are obtained by using the equation stated before. The $\epsilon_{\text{Nd}}(T)$ values also can be obtained in an equivalent manner for the Sm-Nd decay system⁷.

The $\epsilon_{\text{Ce}}(T)$ and $\epsilon_{\text{Nd}}(T)$ values are calculated to be $+1.1$ and -5.9 , respectively, for Broderic Falls granodiorite, $+0.8$ and -5.7 for Bushveld gabbro, $+0.5$ and -8.6 for Morton gneiss and -1.5 and -2.6 for grey tonalitic gneiss, respectively. The $\epsilon(T)$

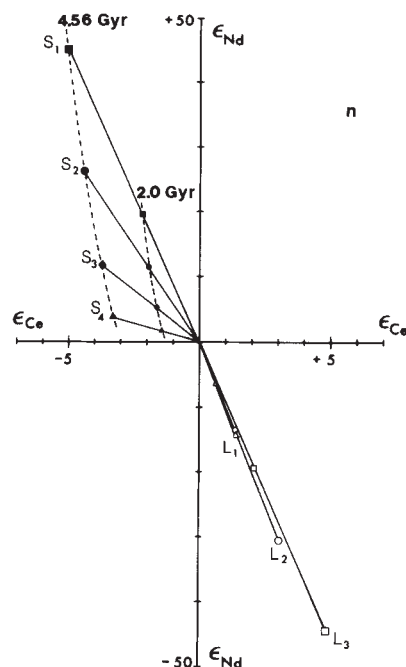


Fig. 3 Theoretically calculated variations of $^{138}\text{Ce}/^{142}\text{Ce}$ and $^{143}\text{Nd}/^{144}\text{Nd}$ values for the rocks with various REE patterns shown in Fig. 2. (For dashed lines, see the text).

values are shown on $\epsilon_{\text{Ce}}-\epsilon_{\text{Nd}}$ diagram together with present $\epsilon(0)$ values (Fig. 4). Initial $\epsilon(T)$ values for Broderic Falls granodiorite, Bushveld gabbro and Morton gneiss are plotted in the same $+\epsilon_{\text{Ce}}$ and $-\epsilon_{\text{Nd}}$ quadrant as present $\epsilon(0)$ values, and extrapolated tie-lines between $\epsilon(T)$ and $\epsilon(0)$ values seem to conjugate at the origin of the diagram. These facts suggest that the source materials of these three rocks have been found only in LREE-rich or chondritic REE features. In Fig. 4, it is seen that the Bushveld gabbro and Broderic Falls granodiorite were derived from source materials with quite similar time-integrated features for La/Ce and Sm/Nd ratios (that is, REE pattern), though the former is gabbro and the latter is felsic granodiorite, and their current REE patterns are different. The calculated $\epsilon_{\text{Nd}}(T)$ value for Morton gneiss seems to be too low for a mantle-derived rock. The 3,000 Myr age chosen here for Morton gneiss may not be the age in indicating derivation from the mantle but an age indicating a remelting event. Neodymium model ages against CHUR are calculated to be 3,800 Myr and 3,700 Myr for Morton gneiss and grey gneiss, respectively. These values agree well with the value (3,600 Myr) of Watersmeet gneiss¹⁴.

The negative ϵ_{Ce} and ϵ_{Nd} initial values of grey gneiss are of interest. These values indicate that source material of the grey gneiss was depleted in La and Sm on the chondrite-normalized pattern as compared with Ce and Nd. This feature suggests a convex-up REE pattern with a summit around praeodymium or neodymium. Some amphibolite and hornblende nodules from Itinome-gata volcano in Japan have a similar REE pattern with a summit around praeodymium or neodymium¹⁵. These nodules are believed to have formed at the lower crust. The grey tonalitic gneiss may be a remelting product of lower crustal materials with a peculiar REE pattern as mentioned above.

Mantle-derived volcanic rocks such as MORB and IAB are rather dispersed in the second quadrant (the $+\epsilon_{\text{Nd}}$ and $-\epsilon_{\text{Ce}}$ area of Fig. 1). The three MORBs are farthest from the CHUR point. The MORB from the Deep Sea Drilling Project (DSDP) Site 396 has a lower ϵ_{Ce} value than the other two MORBs, though the ϵ_{Nd} value is almost identical. The La/Ce ratio of the DSDP Site 396 basalt is indistinguishable from the other two MORBs under consideration. In view of the model calculation (Fig. 3), this DSDP Site 396 basalt is considered to be derived

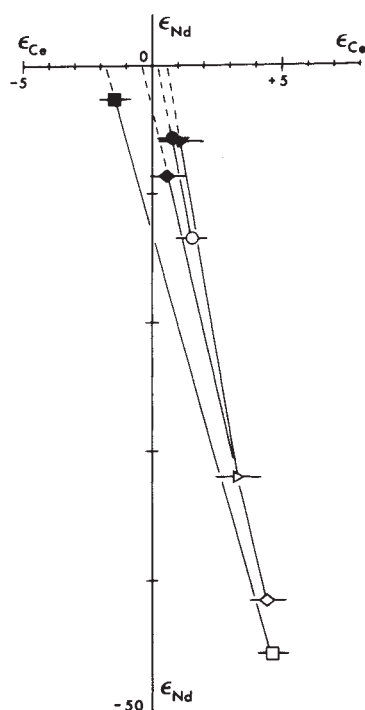


Fig. 4 $\epsilon_{\text{Ce}}-\epsilon_{\text{Nd}}$ diagram of initial isotopic compositions ($\epsilon(T)$) of samples from continental crust (solid symbols), and their present isotopic compositions ($\epsilon(0)$): open symbols). Squares, diamonds, triangles and circles show grey tonalitic gneiss, Morton gneiss, Broderick Falls granodiorite and Bushveld gabbro, respectively. Corresponding marks are tied together with the isotope growth curve (solid line).

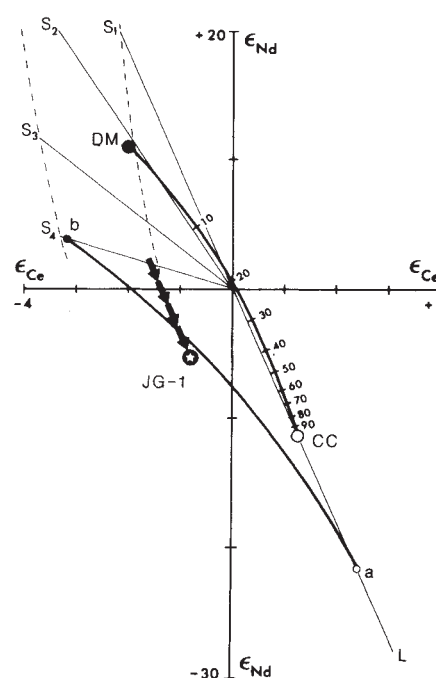


Fig. 5 Isotope growth and mixing features of depleted mantle-derived magma and continental crust. DM and CC indicate depleted mantle and average continental crust, respectively. The line joining a and b is a mixing line, drawn to pass near JG-1. Line L shows $\epsilon_{\text{Ce}} = -0.112 \epsilon_{\text{Nd}}$. Fine solid lines S_1, S_2, S_3, S_4 and dashed lines are the same as Fig. 3.

from a source with a longer history of light REE depletion than the other two MORBs.

IAB and OIB are somewhat closer to the CHUR-point in Fig. 1 than MORB. JR-1 (rhyolite, no. 14 in Table 1) is in the same area as IAB and OIB. This indicates that this rhyolite (an open circle in the $+\epsilon_{\text{Nd}}, -\epsilon_{\text{Ce}}$ quadrant) is not the product of remelting of acidic crust, but is probably a differentiation product of mantle-derived magma with little contamination by crustal REE.

The Columbia River flood basalt (BCR-1) has a $^{138}\text{Ce}/^{142}\text{Ce}$ value (0.0228520) very close to the chondritic value⁹ (0.0228527). This feature is consistent with Nd data: BCR-1 also has the chondritic value for $^{143}\text{Nd}/^{144}\text{Nd}$ (ref. 7). Accordingly, BCR-1 is at the origin in Fig. 1. A Gough ankaramite and an Oki alkali-olivine basalt are also nearly chondritic in both $^{138}\text{Ce}/^{142}\text{Ce}$ and $^{143}\text{Nd}/^{144}\text{Nd}$ ratios (slightly lower than chondritic for Gough ankaramite).

One of the major interests in petrology is the question of whether the rock is the product of simple magmatic differentiation of mantle material or whether the rock has been subjected to any crustal contamination. A hypothetical mixing is examined on the $\epsilon_{\text{Ce}}-\epsilon_{\text{Nd}}$ diagram (Fig. 5). A point labelled DM represents depleted mantle. The point is obtained by an arithmetic average of the three MORB data in Table 1. Goldstein *et al.*¹⁶ estimated the average ϵ_{Nd} value for continental crust as -11.4. A value of +1.3 is obtained as an average ϵ_{Ce} for continental crust by using a well-defined correlation $\epsilon_{\text{Ce}} = -0.112 \epsilon_{\text{Nd}}$ for rocks from continental crust. Curvature of the line of mixing between two components as shown in Fig. 5 depends on the relative concentrations of Nd and Ce for the end members. We assume that the depleted mantle and continent have REE patterns like S_2 and L_2 in Fig. 2, respectively. The absolute value is not an important factor for curvature. A calculated mixing line is shown in Fig. 5. If we assume an REE pattern like L_3 for continental crust instead of L_2 , the mixing curve in Fig. 5 turns out to have a slightly greater convexity. In any case, mixing lines between

continental crust with L-type REE patterns and depleted mantle with S-type REE patterns are expected to have a gently convex-up curve.

The Ce-Nd isotope data of MORB and IAB examined here are rather dispersed in the $+\epsilon_{\text{Nd}}, -\epsilon_{\text{Ce}}$ quadrant as mentioned before. This feature is difficult to explain by the mixing of magma derived from the uniform mantle and materials of continental crust. If there are magma sources within the mantle with various LREE-depleted patterns as represented by S_1, S_2, S_3 and S_4 it is to be expected that the basalts derived from those sources display a $\epsilon_{\text{Nd}}-\epsilon_{\text{Ce}}$ dispersion as observed. According to our theoretical considerations, the observed dispersion reflects the original variations in the REE-characteristics of the source. Comparison of the actual data (Fig. 1) with the model calculations (Fig. 3) leads us to conclude that most of the rocks (except the three rocks plotted around the CHUR-point) were derived from mantle with S_2 - S_3 -type REE patterns, 1,500-2,000 Myr old. Dickin¹⁷ found a Ce-Nd isotope 'mantle array' on eight ocean island basalts. A well-defined mantle array indicates mixing between the thoroughly homogenized mantle and crustal materials¹⁷.

There is a controversy concerning the chondritic ϵ_{Nd} values of BCR-1, as to whether BCR-1 is derived from CHUR⁷ or is a fortuitous result of magma mixing derived from depleted mantle and Precambrian crustal material¹⁸. It is difficult to judge from the Ce-Nd isotopes, which of the above cases is correct, because the mixing line of depleted mantle and crustal materials can easily pass through the origin of Fig. 5.

Japanese granodiorite (JG-1) has a LREE-enriched pattern like the granodiorite from Broderick Falls. However, JG-1 (open circle in the third quadrant in Fig. 1) deviates from the array formed by continental rocks. This suggests that JG-1 is not a direct derivative from a chondritic source. Here is a possible case of the magma derived from a LREE-depleted source mixing with crustal rock. As discussed above, mixing is expected to follow a slightly convex-up curve. To pass the line

near the JG-1 point, the two end members for mixing are required to be quite old. A possible example which would satisfy available constraints is the mixing of a magma derived from S_4 -type source 4,000 Myr old, and 3,000 Myr old crust with L_2 -type REE patterns. However, such old materials cannot be expected in Japanese Islands.

Another way to reach the ϵ_{Ce} and ϵ_{Nd} values of JG-1 is to have a considerable isotope growth in the L-type source which is derived from an S-type primary source. (A production of L type from S type can be realized by a small degree of partial melting of the S type source¹⁹.) As discussed before, isotope growth in L-type material follows a trend of almost constant slope in the ϵ_{Ce} - ϵ_{Nd} diagram (see Fig. 3). If we assume that the L-type source material is derived from an S-type (for example, S_4) primary source, isotope growth is obtained as shown by the arrow in Fig. 5. In this case S_1 -type material cannot be a primary source: this is because the L-type growth line passing through JG-1 does not cross the S_1 growth line at a geologically plausible time. For the petrogenesis of JG-1, of course, crustal contamination cannot be ruled out, but the effect by the contamination might be a secondary contribution, if present at all.

Received 1 September 1986; accepted 13 March 1987.

1. Tanaka, T. & Masuda, A. *Nature* **300**, 515-518 (1982).
2. Nakamura, N., Tatsumoto, M. & Ludwig, K. R. *J. geophys. Res. Suppl.* **89**, B438-B444 (1984).
3. Tanaka, T., Shimizu, H., Shibata, K. & Masuda, A. *Lunar planet. Sci.* **16**, 847-848 (1985).
4. Hamilton, P. J. *J. Petrol.* **18**, 24-52 (1977).
5. Goldich, S. S. & Wooden, J. L. *Spec. Pap. geol. Soc. Am.* **182**, 77-94 (1980).
6. Wasserburg, G. J., Jacobsen, S. B., DePaolo, D. J., McCulloch, M. T. & Wen, T. *Geochim. cosmochim. Acta* **45**, 2311-2323 (1981).
7. DePaolo, D. J. & Wasserburg, G. J. *Geophys. Res. Lett.* **3**, 743-746 (1976).
8. Lugmair, G. W. & Carlson, R. W. *Proc. 9th Lunar Planet. Sci. Conf.* 689-704 (1978).

Conclusion

Within the Nd-Sm isotope diagram there is a remarkably well-defined 'mantle array', which has provided tremendous insight into several important geochemical and geological questions. It is not yet clear whether there is a similar clearly defined mantle array in the Ce-Nd isotope system or not. Further improvement of analytical techniques, and accumulation of more data on cerium isotopes should enable us to investigate the evolution and the scale of mantle heterogeneities in more detail. In contrast to the present status of mantle Ce-Nd data, the Ce-Nd array for rocks of the continental crust is very well-defined. This Ce-Nd array is a very distinctive characteristic of continental crustal materials. La-Ce isotope systematics may make a large contribution to studies on crustal development.

We thank Drs P. Taylor and W. White for constructive suggestions, and Drs S. S. Goldich, J. Keith, K. Suwa, K. Shibata, S. Aramaki, I. Kaneoka, T. Ui and T. Fujii for samples of Minnesota, Bushveld, Broderick Falls, Deccan, Gough, Hawaii and the Mid-Atlantic Ridge, respectively.

9. Shimizu, H., Tanaka, T. & Masuda, A. *Nature* **307**, 251-252 (1984).
10. O'Nions, R. K. & Pankhurst, R. J. *Earth planet. Sci. Lett.* **22**, 328-338 (1974).
11. Arth, J. G. & Hanson, G. N. *Contr. Miner. Petrol.* **37**, 161-174 (1972).
12. Haskin, L. A., Haskin, M. A., Frey, F. A. & Wildeman, T. R. in *Origin and Distribution of the Elements* (ed. Ahrens, L. H.) 889-912 (Pergamon, Oxford, 1968).
13. Masuda, A. & Jibiki, H. *Geochim. J.* **7**, 55-65 (1973).
14. McCulloch, M. T. & Wasserburg, G. J. *Spec. Pap. geol. Soc. Am.* **182**, 135-138 (1980).
15. Tanaka, T. & Aoki, K. *J. Geol.* **89**, 369-390 (1981).
16. Goldstein, S. L., O'Nions, R. K. & Hamilton, P. J. *Earth planet. Sci. Lett.* **70**, 221-236 (1984).
17. Dickinson, A. P. *Nature* (in the press).
18. Carlson, R. W., Lugmair, G. W. & Macdougall, J. D. *Nature* **289**, 160-162 (1981).
19. Masuda, A. & Matsui, Y. *Geochim. cosmochim. Acta* **30**, 239-250 (1966).

Highly potent and selective ligands for histamine H_3 -receptors

J.-M. Arrang*, M. Garbarg*, J.-C. Lancelot†, J.-M. Lecomte‡, H. Pollard*, M. Robba†, W. Schunack§ & J.-C. Schwartz*

* Unité 109 de Neurobiologie et Pharmacologie, Centre Paul Broca de l'INSERM, 2ter rue d'Alésia, 75014 Paris, France

† Laboratoire de Pharmacie Chimique, UER des Sciences Pharmaceutiques, 1 rue Vaubénard, 14032 Caen, France

‡ Laboratoire Bioprojet, 30 rue des Francs Bourgeois, 75003 Paris, France

§ Institut für Pharmazie, Freie Universität Berlin, Konigin-Luise-Strasse 2+4, D-1000 Berlin, FRG

New drugs selective for histamine H_3 -receptors can be used to establish that these receptors are involved in the feedback control of histamine synthesis and release, and to demonstrate their distribution in the brain and peripheral tissues. These drugs provide new tools for affecting physiological and possibly pathological conditions in which histamine is involved.

SINCE the discovery of histamine by Sir Henry Dale at the beginning of this century, the multiple functions of this amine as a chemical messenger in cell-to-cell communication have been progressively revealed. Histamine is stored in, and released from, various cellular populations including immunoglobulin E (IgE) receptor-bearing cells (mast-cells and basophils)¹, endocrine cells² and neurons in the central³ or peripheral nervous system⁴. An even larger variety of cell types including smooth muscles, neurons, endocrine or exocrine glands, blood cells and cells of the immune system, respond to histamine by increasing the intracellular levels of signals generated by either the phosphatidylinositol cycle⁵ or adenylate cyclase⁶. Each of these two intracellular responses is mediated by a pharmacologically distinct subclass of receptors, the H_1 - and H_2 -receptors respectively⁷.

More recently it was observed that histamine inhibits its own synthesis and release in brain slices by a negative feedback process operating at the level of histaminergic nerve-endings⁸⁻¹⁰.

It was suggested that these actions of histamine are mediated by a novel subclass of receptors (H_3) because they are not blocked by H_1 -receptor antagonists and, although some H_2 -receptor antagonists were effective, their potencies differed from those found in reference H_2 -receptor-mediated responses. No selective agonist or antagonist of histamine could be identified at putative H_3 -receptors. Because a receptor subclass is ultimately defined by the use of highly selective agents and its function assessed by testing the agents in the living animal, we have undertaken to identify or design suitable compounds. We now describe the actions of (*R*)- α -methylhistamine (α -MeHA), a chiral agonist, and thioperamide, an antagonist derived from imidazolyloperidine, which both display high selectivity and potency at nanomolar concentrations *in vitro*. These drugs modify antagonistically histamine synthesis and/or release from brain slices and also cerebral and peripheral tissues in the living rat. In addition a partial agonist at H_3 -receptors is described. Finally we show that [³H]-(*R*)- α -MeHA is a useful probe for

Table 1 Potencies of various imidazole derivatives at the three subclasses of histamine receptors

Agents	H ₁ -Receptor	H ₂ -Receptor	H ₃ -Receptor			
			Inhibition of [³ H]histamine release	Inhibition of [³ H]histamine synthesis	Inhibition of [³ H]-(R) α -MeHA binding Rat brain	Inhibition of [³ H]-(R) α -MeHA binding Guinea-pig lung
Agonists (relative potencies)						
Histamine	100	100	100	100	100	100
(<i>R</i>)- α -MeHA	0.49*	1.02‡	1,550	1,130	1,060	813
(<i>S</i>)- α -MeHA	0.49*	1.74‡	13		11	52
4-[2-(1-Pyrrolidinyl)ethyl]-imidazole	Antagonist* (K _i = 5 μ M)	0.1‡ (30% max.)	7 (56% max.)		5	6
Antagonist (K _i values)						
Thioperamide	>100 μ M†	>10 μ M§	4.3 nM	31 nM	2.1 nM	2.0 nM

The relative potencies were calculated as the ratio (EC₅₀ (or K_i in binding experiments) of histamine)/(EC₅₀ (or K_i) of agonist) $\times$ 100. The EC₅₀ and K_i were determined from data shown in Figs 1 and 2. The relative potency of (*R*)- α -MeHA in the inhibition of [³H]-(*R*)- α -MeHA binding in rat brain was calculated from its K_D. The K_i of thioperamide in relation to the inhibition of [³H]histamine release is the mean value of its various determinations described in Fig. 1 legend.

* Tested for stimulating contraction of the isolated guinea-pig ileum. Data from refs 44 and 45.

† Tested for competition with [³H]mepyramine binding in membranes of guinea-pig cerebellum.

‡ Tested for rate stimulation in the spontaneously beating isolated guinea-pig right atrium. Data from ref. 44.

§ Tested for inhibition of cyclic AMP accumulation in slices from guinea-pig hippocampus.

the radioassay and autoradiographic visualization of H₃-receptors.

A chiral agonist

Histamine is an achiral molecule which presumably adopts different preferred conformations when interacting productively with its various receptor subclasses. It appears to act at both H₁- and H₂-receptors as the *N*⁷-tautomer of the monocation, possibly adopting the *trans* conformation¹¹. Analysis of the chemical properties of selective H₁- or H₂-receptor agonists have provided further insights into the molecular requirements of the corresponding receptor subclasses¹¹. But none of these compounds is able to mimic the actions of histamine at H₃-receptors and some of them even behave as antagonists⁸⁻¹⁰. In addition a series of chiral histamine derivatives with branched side chains were used to elucidate the stereochemical requirements of the various receptor subclasses¹². With these compounds no stereo selectivity and limited stereoselectivity were found at H₁- and H₂-receptors respectively, whereas a high degree of stereoselectivity, the reverse of that observed at H₂-receptors, was recently found for two compounds (*N* ^{α} -methyl- α -chloromethyl-histamine and α , *N* ^{α} -dimethylhistamine) at H₃-receptors. However both active enantiomers behaved as weak agonists, their potency at H₃-autoreceptors being only 1–4% that of histamine¹³.

In contrast, by systematically studying chiral analogues with a single substitution of the side chain in an attempt to reduce the conformational freedom of the histamine molecule and thereby increase its specificity without losing its activity, we have identified α -MeHA as a compound that fulfills these requirements. In the H₃-receptor assay system in which depolarization-induced release of [³H]histamine from brain slices is evaluated⁸, both (*R*) and (*S*)- α -MeHA progressively inhibited the release by up to about 60% that is, to the same maximal extent as exogenous histamine itself (Fig. 1a). This suggests that both enantiomers behaved as full H₃-receptor agonists, but the (*R*) isomer exerted this effect at nanomolar concentrations, being approximately 100 times as potent as the (*S*) isomer and 15 times as potent as histamine itself. This is in agreement with previous observations showing that enantiomers corresponding to *S*-configured L-histidine are highly preferred at H₃-receptors whereas enantiomers corresponding to *D*-histidine are slightly more potent at H₂-receptors, no difference being observed at H₁-receptors¹³. As the (*R*)- α -MeHA potencies relative to histamine are only 0.5% and 1% at reference H₁- and H₂-receptor systems respectively (Table 1), it can be regar-

ded as a highly potent and selective H₃-receptor agonist. Because H₃-receptors are more sensitive to histamine and agonists than H₁- or H₂-receptors, (*R*) α -MeHA might even be regarded as more selective if one considers its *pD*₂ value (negative logarithm of the EC₅₀), which is, 4.54, 3.96 and 8.40 at the reference H₁-, H₂- and H₃-receptor systems respectively. This means that stimulation of H₃-receptors by (*R*)- α -MeHA is expected to occur at concentrations $\sim$ 10,000 times lower than those required for H₁- or H₂-receptor stimulation.

The H₃-autoreceptors seem to control not only release of [³H]histamine from, but also its synthesis in, depolarized cerebral histaminergic neurons^{10,14}. Consistently, (*R*)- α -MeHA potentially inhibited the K⁺-induced stimulation of [³H] histamine formation in slices of rat cerebral cortex (Fig. 1d).

That these effects of (*R*)- α -MeHA are mediated by H₃-receptors was confirmed by their blockade by impromidine⁸ (not shown) and the novel antagonist thioperamide (see below).

A competitive antagonist

In seeking a specific antagonist of histamine at cerebral H₃-receptors able to cross the 'blood-brain barrier', the possibility of using known H₁- or H₂-receptor ligands as chemical starting point was first considered. But the lipophilic H₁-receptor antagonists tested displayed negligible affinity⁸. In contrast several agonists at H₁-receptors, such as betahistidine¹⁵, or at H₂-receptors, such as impromidine, as well as H₂-receptor antagonists, such as burimamide, displayed significant antagonist activity⁸. But these are generally highly hydrophilic and/or positively charged molecules with little ability to enter the brain from the circulation.

Hence we took a completely different approach in which the histamine molecule was used as the chemical starting point and analogues were synthesized in which the ethylamine side chain was extended and its nitrogen atom included in a piperidine ring. In all cases this resulted in a loss of agonist activity. The affinity of compounds for H₃-receptors was progressively optimized by varying the nature of the piperidine-substituting groups and thioperamide (MR 12842, *N*-cyclohexy-4-(imidazol-4-yl)-1-piperidinecarbothioamide), was finally selected.

When thioperamide was opposed at a fixed concentration to histamine, the concentration-response curve to the amine was shifted to the right without change of the maximal response, leading¹⁶ to an apparent K_i of 7.6 nM (Fig. 1a). Similar apparent dissociation constants for thioperamide were derived¹⁷ from the analysis of the progressive reversal of the inhibition of [³H]histamine release exerted by either exogenous histamine or

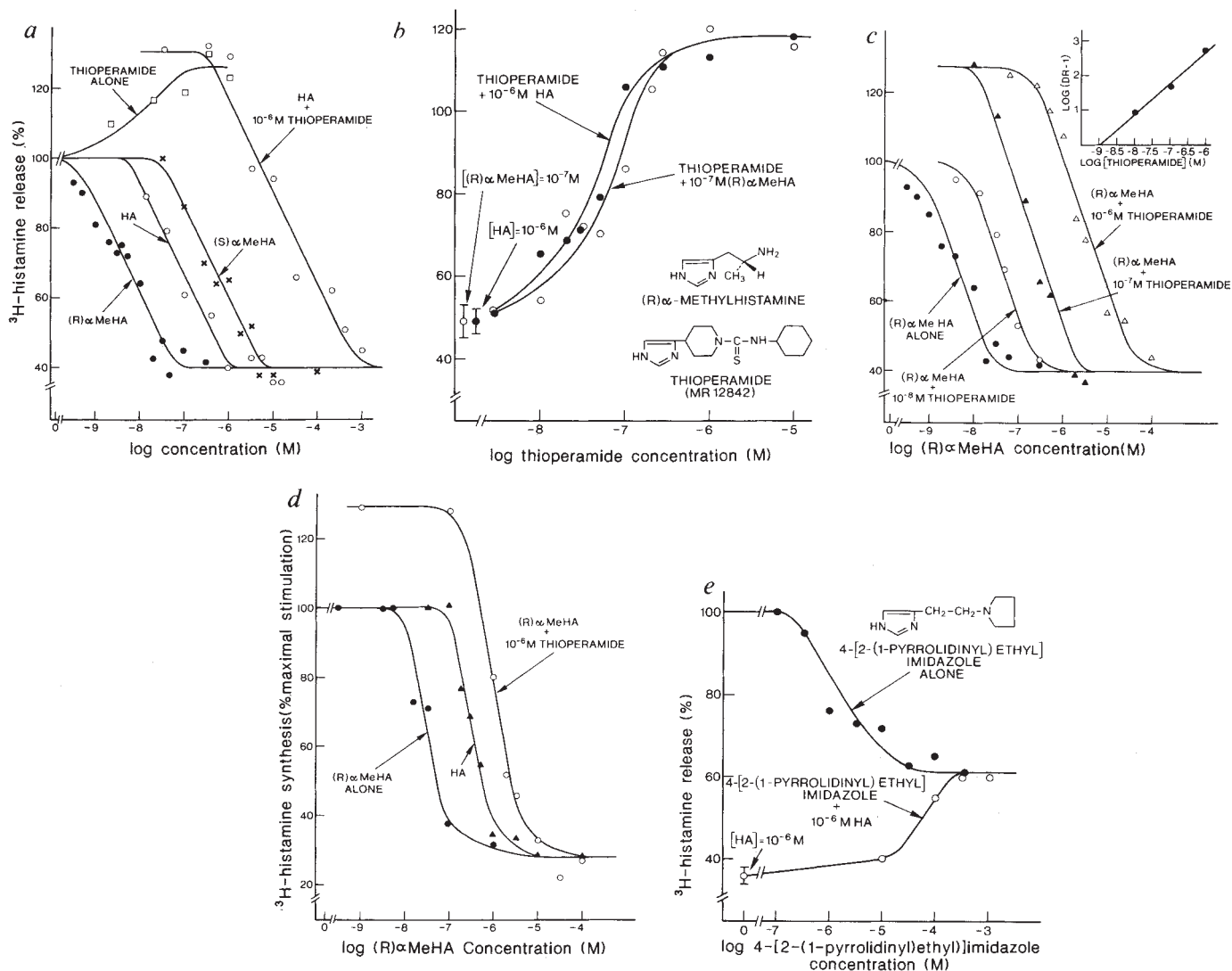


Fig. 1 Effects of various imidazole derivatives on the K^+ -induced release of [3H]histamine from, and stimulation of [3H]histamine synthesis in, slices of rat cerebral cortex. **a**, Release of [3H]histamine was inhibited by histamine (HA) ($\circ$, left-hand slope), (R)- α -MeHA ($\bullet$) and (S)- α -MeHA ($\times$) with EC_{50} values (nM) of 62 ± 14 , 4.0 ± 0.8 and 470 ± 160 respectively. Maximal inhibitions were $62 \pm 2\%$ in all cases. Thioperamide alone ($\square$) maximally increased [3H]histamine release by $24 \pm 3\%$ (half-maximal concentration: 4.1 ± 3.5 nM). In the presence of $1 \mu M$ thioperamide the EC_{50} of histamine became $8.2 \pm 2.6 \mu M$, leading to a K_i of 7.6 nM for the antagonist ($\circ$, right hand slope). **b**, The IC_{50} values (nM) of thioperamide were 48 ± 9 and 74 ± 17 when opposed to histamine (HA, $\bullet$) and (R)- α -MeHA ($\circ$) respectively, leading to K_i values (nM) of 5.6 ± 1.0 and 2.9 ± 0.6 . **c**, Thioperamide at fixed concentrations was opposed to (R)- α -MeHA in increasing concentrations (release experiments) ($\bullet$, (R)- α -MeHA alone; and with $\circ$, 10^{-8} M; $\blacktriangle$, 10^{-7} M and $\triangle$, 10^{-6} M thioperamide). Inset, Schild plot of data in which dose ratios (DR) were determined from EC_{50} values of (R)- α -MeHA. The slope and 95% confidence limits for the regression of $\log (DR-1)$ on $\log [thioperamide]$ was 0.91 ($0.82-1.01$), the coefficient of correlation was 0.995 and the pA_2 8.96 . **d**, The K^+ -induced stimulation of [3H]histamine synthesis was inhibited ($-71 \pm 4\%$) by histamine and (R)- α -MeHA with EC_{50} values (nM) of 340 ± 30 and 30 ± 7 respectively. Symbols: $\bullet$, (R)- α -MeHA alone; $\blacktriangle$, histamine; $\circ$, (R)- α -MeHA with 10^{-6} M thioperamide. **e**, 4-[2-(1-pyrrolidinyl)ethyl]imidazole alone ($\bullet$) inhibited the K^+ -evoked release of [3H]histamine by $36 \pm 2\%$ (as compared to $64 \pm 2\%$ for histamine) and with an EC_{50} of $0.9 \pm 0.3 \mu M$. When opposed to histamine ($\circ$) its IC_{50} was $38 \pm 7 \mu M$ and a plateau corresponding to $39 \pm 1\%$ inhibition of release was reached. From this an intrinsic activity of 56% that of histamine and a K_i of $2.2 \pm 0.4 \mu M$ were derived.

Methods. In **a**, **b**, **c** and **e**, the K^+ -evoked release of [3H]histamine was evaluated as described⁸. After preincubation with L-[3H]histidine (Amersham), slices were incubated for 2 min with 2 or 30 mM K^+ . Drugs were added 5 min before the start of incubations, at the end of which [3H]histamine was assayed⁴⁷ in tissue and medium. The spontaneous efflux of [3H]histamine into the medium (2 mM K^+) represented $5.0 \pm 0.2\%$ of the total (tissue plus medium) and was not altered by any of the added agents. In the absence of added agents, the [3H]histamine release evoked by 30 mM K^+ (expressed as percentage of total [3H]histamine released in the medium over spontaneous efflux) represented $14.2 \pm 0.6\%$. Results are expressed as percentages of this value. Means from 2-5 different experiments with at least triplicate determinations. In **d**, the K^+ -induced stimulation of [3H]histamine synthesis was evaluated in slices prepared as described⁸. Aliquots of the slice suspension (~ 4 mg protein) were incubated for 30 min at $37^\circ C$ under a stream of O_2/CO_2 (95:5) in the presence of $0.55 \mu M$ [3H]histidine and 2 or 30 mM K^+ . Drugs were added at the beginning of incubations which were stopped by sonication of the total samples (tissue plus medium) and addition of $HClO_4$ (0.4 M final concentration). The [3H]histamine was isolated on two successive Amberlite CG 50 columns⁴⁷ with a $77 \pm 3\%$ recovery, contamination by [3H]histidine being $<0.001\%$. In the absence of added agents [3H]histamine synthesis represented (d.p.m. per mg protein) $2,055 \pm 124$ (2 mM K^+) and $3,953 \pm 248$ (30 mM K^+). The difference between these two values was taken as the K^+ -induced stimulation and represented $1,898 \pm 277$ d.p.m. per mg protein (92% increase). Results are expressed as percentages of this value and correspond to 2-7 experiments with at least triplicate determinations. For determination of EC_{50} and IC_{50} values, the total curves were analysed with an iterative computer least-squares method⁴⁹. Values of K_i were determined neglecting the influence of endogenous histamine.

(R)- α -MeHA in fixed concentration (Fig. 1b). Also, Schild plot analysis of the progressive shift to the right of the concentration-response curve to (R)- α -MeHA elicited by thioperamide, compatible with a competitive antagonism by this compound, led to a pA_2 of 8.96 , corresponding to a K_i of 1.1 nM (Fig. 1c). In addition, thioperamide alone elicited in a concentration-depen-

dent manner a small but consistent increase ($\sim 25\%$) in [3H]histamine release with half-maximal effect at 4.1 nM (Fig. 1a). This facilitatory effect was previously observed with other H_3 -receptor antagonists and attributed to antagonism of endogenous histamine released by the depolarizing stimulus and exerting its normal negative feedback influence on the release

process. All these consistent data indicate that thioperamide competitively blocks H_3 -autoreceptors regulating [3H]histamine release with a mean apparent K_i of 4 nM, ignoring the influence of endogenous histamine (Table 1). The competitive nature of thioperamide antagonism was confirmed in radioligand binding studies.

Finally, thioperamide also antagonized in a surmountable manner the (R)- α -MeHA-induced inhibition of [3H]histamine synthesis with a K_i of 31 nM (Fig. 1d). This value, slightly higher than that found in the release studies, may conceivably result from a stronger influence of endogenous histamine, reaching higher extracellular concentrations after the longer-lasting depolarizations used in synthesis studies. Thioperamide has negligible affinity for H_1 - and H_2 -receptors (Table 1) as well as for α - or β -noradrenaline^{18,19} ($K_i > 1 \mu M$), muscarinic²⁰ ($K_i > 10 \mu M$), opiate²¹, dopamine²² or serotonin^{23,24} ($K_i > 0.1 mM$) receptors. This indicates that thioperamide is a potent and selective H_3 -receptor antagonist.

A partial agonist

Although alkylation of the histamine side-chain terminal nitrogen with bulky groups leads to a loss of H_3 -receptor agonist activity²⁵, its introduction into a pyrrolidine nucleus led to a compound which was still able to mimic histamine partially in inhibiting the [3H]amine release (Fig. 1e). However, the relative potency was reduced to 7% and, interestingly, the maximal response was only 56% that of histamine, suggesting that this compound may represent a partial agonist. This was confirmed by opposing it to histamine, whose maximal response was reduced to the same level, leading to a K_i of 2.2 μM for 4-[2-(1-pyrrolidinyl)ethyl] imidazole when considered as a pure antagonist.

Action in vivo

The availability of potent agents acting selectively at H_3 -receptors made it possible for the first time to assess the physiological function of these receptors in the control of histamine synthesis and release in brain and peripheral organs of living animals. In rats treated with thioperamide (5 mg per kg, intraperitoneally (i.p.)) the steady state histamine level in a subcellular fraction from cerebral cortex enriched in nerve endings was reduced by 30% after 30 min (Table 2). This presumably reflected an increased rate of [3H]histamine release because the H_3 -receptor antagonist also accelerated by about twofold the amine depletion observed after irreversible inhibition of its synthesis by α -fluoromethylhistidine²⁶. The reversal of this effect by the agonist (R)- α -MeHA (20 mg per kg, i.p.) is consistent with the idea that it is H_3 -receptor-mediated (Table 2). That (R)- α -MeHA alone modestly increased the steady state level of histamine but completely prevented the histamine depletion elicited by inhibition of its synthesis presumably reflects a marked reduction of histamine release *in vivo*. This effect of the agonist was stereoselective, as shown by the non-significant change in α -fluoromethylhistidine-induced depletion 45 min after administration of (S)- α -MeHA which, *in vitro*, is 100-fold less potent than the (R) isomer.

Such opposite effects of the agonist and the antagonist were also found when [3H]histamine formation in cerebral cortex was evaluated after administration of its 3H -labelled precursor. Hence (R)- α -MeHA (10 mg per kg, intravenously (i.v.)) significantly decreased [3H]histamine formation by about 30% and this was reversed by thioperamide (10 mg per kg, i.p.) which, alone or in combination with (R)- α -MeHA, markedly enhanced the 3H -labelled amine formation (Table 3). A qualitatively similar effect occurred in the hypothalamus, where histamine perikarya (and also terminals) are located^{3,27,28}. Hence it seems that both α -MeHA and thioperamide cross the blood-brain barrier at reasonable dosages and that brain histamine neurons activity can be modulated bi-directionally by occupation of

Table 2 Effects of α -methylhistamine and thioperamide on steady-state level and metabolism of histamine in rat cerebral cortex

Treatment	Time after treatment (min)	Histamine level (ng per g protein)	(% change)
Control		547 $\pm$ 24	
Thioperamide	30	381 $\pm$ 21†	-30
(R)- α -MeHA	30	626 $\pm$ 47	+15
(S)- α -fluoromethylhistidine:			
alone	30	438 $\pm$ 30*	-20
+thioperamide	30	313 $\pm$ 11†‡	-43
+(R)- α -MeHA	30	653 $\pm$ 38*‡	+19
+thioperamide	30	602 $\pm$ 56‡	+10
+(R)- α -MeHA			
Control		532 $\pm$ 26	
(S)- α -Fluoromethylhistidine:			
alone	45	305 $\pm$ 29†	-43
+(R)- α -MeHA	45	605 $\pm$ 23‡	+14
+(S)- α -MeHA	45	367 $\pm$ 17†	-31

Male Wistar-rats (80–110 g) received the various drugs i.p. and were killed by decapitation 30 or 45 min later. Doses of (R)- and (S)- α -MeHA hydrochloride were given at 20 mg per kg (expressed as the base), thioperamide at 5 mg per kg and (S)- α -fluoromethylhistidine, an irreversible inhibitor of L-histidine decarboxylase⁴⁶ (from Dr J. Kollonitsch, MSD Research Laboratories) at 100 mg per kg. The cerebral cortex was dissected out in the cold, homogenized in 10 vols of ice-cold 0.32 M sucrose using a Teflon-glass Potter-Elvehjem homogenizer (clearance 0.10–0.15 mm). A crude P_2 synaptosomal fraction was prepared by two centrifugations (10,000 and 400,000g min respectively), the last pellet being resuspended in phosphate buffer (K_2/K , 20 mM, pH 7.6). Endogenous histamine level in this fraction was estimated with a radioenzymatic assay⁴⁷ using [3H]S-adenosyl methionine (76 Ci mmol⁻¹; Amersham) and a purified preparation of rat kidney histamine N-methyltransferase. The inactivation of L-histidine decarboxylase in the synaptosomal fraction from (S)- α -fluoromethylhistidine-treated rats was controlled using a radioassay⁴⁷. Inhibition was nearly total (90 $\pm$ 2%) from 15 min to 240 min. Histamine depletion in rats receiving (S)- α -fluoromethylhistidine alone had an initial half-life of about 40 min and the maximal decrease was 62% after 120 min (not shown), consistent with previous reports^{3,48}. Values are means $\pm$ s.e.m. from groups of 8–12 treated rats and 8–21 vehicle-treated rats (controls) respectively.

* $P < 0.006$ and † $P < 0.0001$ as compared with controls. ‡ $P < 0.02$ as compared with (S)- α -fluoromethylhistidine alone.

Table 3 Effects of (R)- α -methylhistamine and thioperamide on [3H]histamine formation in various rat tissues

Tissues	[3H]histamine levels (d.p.m. g ⁻¹ $\times 10^{-2}$)			
	Controls	(R)- α -MeHA	Thioperamide	(R)- α -MeHA + thioperamide
Cerebral cortex	32 $\pm$ 2	22 $\pm$ 2† (-29%)	75 $\pm$ 8† (+138%)	63 $\pm$ 5† (+99%)
Hypothalamus	103 $\pm$ 6	88 $\pm$ 5 (-15%)	253 $\pm$ 28† (+145%)	212 $\pm$ 22† (+105%)
Lung	51 $\pm$ 4	37 $\pm$ 3† (-29%)	96 $\pm$ 16† (+86%)	73 $\pm$ 12* (+42%)
Abdominal skin	144 $\pm$ 17	94 $\pm$ 13* (-34%)	109 $\pm$ 18	106 $\pm$ 33
Spleen	26 $\pm$ 3	18 $\pm$ 2* (-31%)	32 $\pm$ 6	25 $\pm$ 5
Colon	61 $\pm$ 5	50 $\pm$ 5	49 $\pm$ 8	46 $\pm$ 5

Groups of male Wistar rats (80–110 g) received 250 μCi of previously purified [3H]L-histidine in the tail vein 10 min before being killed. The (R)- α -MeHA was administered i.v. at 10 mg per kg together with [3H]L-histidine, whereas thioperamide (10 mg per kg, i.p.) was administered 1 h before. The various tissues of each rat were rapidly dissected out and homogenized in 10 vol (w/v) of ice-cold 0.4 M HClO₄. The resulting homogenates were centrifuged (50,000g min) and [3H]histamine was isolated from the supernatants by chromatography on two successive Amberlite CG-50 columns, then quantified by liquid scintillation spectrometry⁴⁷. The recovery of [3H]histamine was 73% whereas contamination by [3H]L-histidine was 0.016% and results were corrected accordingly. The various treatments did not significantly modify the total radioactivity in tissues. Values are means $\pm$ s.e.m. of measurements from 13–38 rats.

* $P < 0.04$; † $P < 0.005$ as compared with controls.

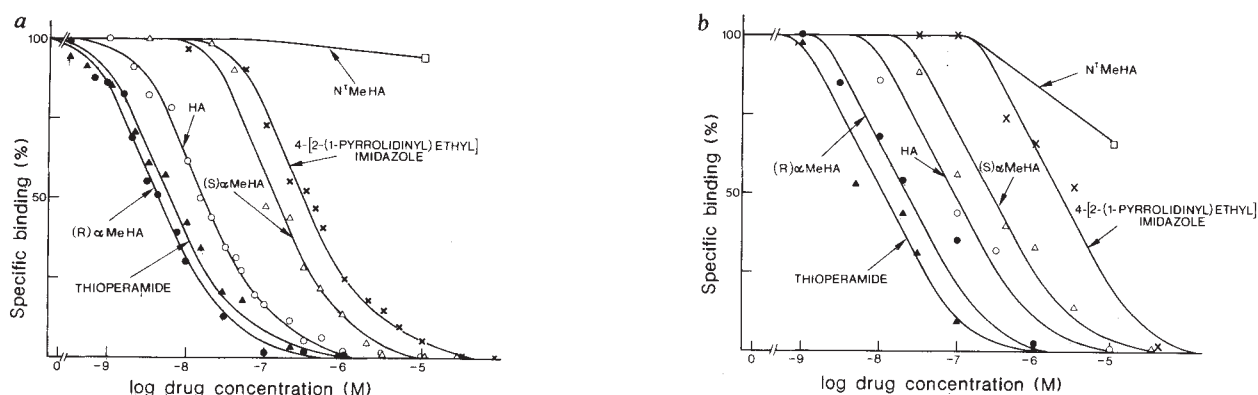


Fig. 2 Inhibition of $[^3\text{H}](R)\alpha\text{-MeHA}$ binding to membranes of rat cerebral cortex (a) and guinea-pig lung parenchyma (b) by various imidazole derivatives. a, Values of K_i for the compounds were determined from their IC_{50} according to the equation¹⁷ $K_i = \text{IC}_{50}/(1 + S/K_D)$ where S is the $[^3\text{H}](R)\alpha\text{-MeHA}$ concentration (1 nM) and K_D its dissociation constant (0.5 ± 0.2 nM) as determined from saturation kinetics at equilibrium. Pseudo Hill coefficients of these compounds did not significantly differ from unity. Values of K_i (nM) were 5.3 ± 0.3 (histamine $\circ$), 1.3 ± 0.1 ($(R)\alpha\text{-MeHA}$, $\bullet$), 50 ± 10 ($(S)\alpha\text{-MeHA}$, Δ), 110 ± 10 (4-[2-(1-pyrrolidinyl)ethyl]imidazole, $\times$), 2.1 ± 0.3 (thioperamide, $\blacktriangle$), $>3,000$ ($N^{\text{r}}\text{MeHA}$, the major histamine metabolite in brain, $\square$), and are reported as such or as relative potencies for agonists in Table 1. The K_i of ten additional compounds (including impromidine, burimamide, cimetidine, tiotidine, ranitidine, mepyramine) were highly correlated ($r = 0.95$) with their K_i at H_3 -receptors controlling $[^3\text{H}]$ histamine release (J.-M.A., M.G., J.-L. Morgat, H.P., J. Roy, W.S. and J.-C.S., manuscript in preparation). b, Values of K_i , calculated as described in a and assuming the same K_D for $[^3\text{H}](R)\alpha\text{-MeHA}$, were (nM) 26 ± 40 (histamine), 3.2 ± 2.8 ($(R)\alpha\text{-MeHA}$), 50 ± 16 ($(S)\alpha\text{-MeHA}$), 460 ± 120 (4-[2-(1-pyrrolidinyl)ethyl]imidazole, $\times$), 2.0 ± 0.7 (thioperamide) and $>2,000$ ($N^{\text{r}}\text{MeHA}$). Symbols, as in a. In addition, K_i values for mepyramine and tiotidine were $>1 \mu\text{M}$ (not shown).

Methods. In a, a particulate fraction was obtained after homogenization of the cerebral cortex in 50 volumes (w/v) of ice-cold Na_2/K phosphate buffer 50 mM, pH 7.5 and two successive centrifugations (1,000g min and 200,000g min). After resuspension of the last pellet in fresh buffer, 0.9 ml aliquots of the suspension (0.4 mg protein) were incubated for 60 min at 25°C with 1 nM of $[2.5\text{-}^3\text{H}](R)\alpha\text{-MeHA}$ (specific activity 20 Ci mmol^{-1}) and unlabelled substances in a final volume of 1 ml. After addition of 6 ml of ice-cold buffer, the mixture was filtered over vacuum onto Millipore AAWP filters which were then rinsed twice with 5 ml ice-cold buffer. Specific binding was defined as that inhibited by $1 \mu\text{M}$ thioperamide and represented 250 d.p.m. out of a total of 400 d.p.m. ($62 \pm 2\%$). Results are expressed as percentages of this value, each point representing the results from one to three different experiments with triplicate determinations. With all agents maximal inhibition did not exceed specific binding. In b, guinea-pigs (300–500 g) were killed and the heart and lungs removed together and placed on ice. The airways were carefully combed to separate parenchymal tissue from the bronchial tree. The lung parenchyma was homogenized in 10 volumes (w/v) of ice-cold phosphate buffer and submitted to two successive centrifugations (5,000 and 550,000g min respectively). In order to eliminate extensively endogenous histamine from the preparation, the pellet was washed six times by resuspension in fresh buffer and recentrifugation at 550,000g min. The final pellet was resuspended in fresh phosphate buffer and 0.9 ml aliquots (0.5 mg protein) were incubated as in a, except that $[^3\text{H}](R)\alpha\text{-MeHA}$ concentration was 2.0 nM and filters were Millipore SMWP. Radioactivity of the filters was counted in ACS scintillation fluid at 40% efficiency and 98% reproducibility. At least 2,000 disintegrations were counted over a background of 16 c.p.m. Specific binding, defined as in a, represented 50 d.p.m. out of a total of 160 d.p.m. ($31 \pm 5\%$). Results are expressed as percentages of this value, each point representing the results from one to three different experiments with at least triplicate determinations. With all agents maximal inhibition did not exceed specific binding.

H_3 -receptors. Thioperamide seems to be the first drug known clearly to increase this activity.

Most interestingly, $[^3\text{H}]$ histamine synthesis was modified in lung in the same manner as in brain (Table 3), suggesting that there are H_3 -receptors there as well, a hypothesis confirmed by radioligand binding studies. Because a large part of lung histamine is located in mast cells found in abundance in the bronchial mucosa, in the alveolar septal connective tissue and in the pleura²⁹, it seems likely that these cells are endowed with H_3 -receptors controlling the synthesis of the amine and, possibly, its release.

The effects of the H_3 -receptor antagonist on $[^3\text{H}]$ histamine formation were less clearcut in other peripheral tissues. For instance in skin, where mast cells of the connective tissue type are abundant³⁰, or in spleen, $(R)\alpha\text{-MeHA}$ also decreased $[^3\text{H}]$ histamine formation by about 30% but the reversal by thioperamide was less clear because the H_3 -receptor antagonist did not promote $[^3\text{H}]$ histamine formation over control levels. In these tissues, although the presence of H_3 -receptors remains to be confirmed by other approaches, it cannot be excluded that the modest effects of thioperamide are related to the absence of a tonic release of histamine under basal conditions.

Receptor assay and visualization

The high apparent affinity of $(R)\alpha\text{-MeHA}$ in functional studies suggested that the drug might constitute, when radiolabelled, a suitable probe for assay and visualization of H_3 -receptors.

Indeed $[^3\text{H}](R)\alpha\text{-MeHA}$ was found to bind in a reversible

and saturable manner to cerebral cortex membranes with a K_D of 0.5 nM (derived from either saturation kinetics or dissociation/association rates) and a maximal number of sites representing 30 ± 3 fmol per mg protein (J.-M.A., M.G., J.-L. Morgat, H.P., J. Roy, W.S. and J.-C.S., manuscript in preparation). In comparison H_1 - and H_2 -receptors appear significantly more abundant^{31–33}. Sites labelled with $[^3\text{H}](R)\alpha\text{-MeHA}$ were pharmacologically identified as H_3 -receptors by competition studies with a variety of compounds (Fig. 2a) whose potencies relative to histamine (for agonists) or K_i values (for antagonists) were in good agreement with corresponding values derived from functional studies (Table 1). Nevertheless, note that K_i values of histamine and agonists in binding studies were generally 5–10 times lower than their corresponding EC_{50} values in functional studies. This difference might correspond to differences in the ionic composition of media used in binding and functional studies or to the selective labelling of H_3 -receptors in a discrete conformational state with high affinity for agonists.

In membranes of guinea-pig lung, specific binding sites could also be detected and unambiguously characterized as H_3 -receptors (Fig. 2b and Table 1) in spite of their extremely low density (5 ± 2 fmol per mg protein), which confirms the presence of H_3 -receptors outside the brain. In autoradiographic studies performed with $[^3\text{H}](R)\alpha\text{-MeHA}$ (Fig. 3), H_3 -receptors were found to be fairly widespread in rat brain as previously shown in release studies⁹. Telencephalic areas such as the cerebral cortex (particularly in its most rostral part), striatum, hippocampus (molecular layer of the dentate gyrus), lateral septum, bed

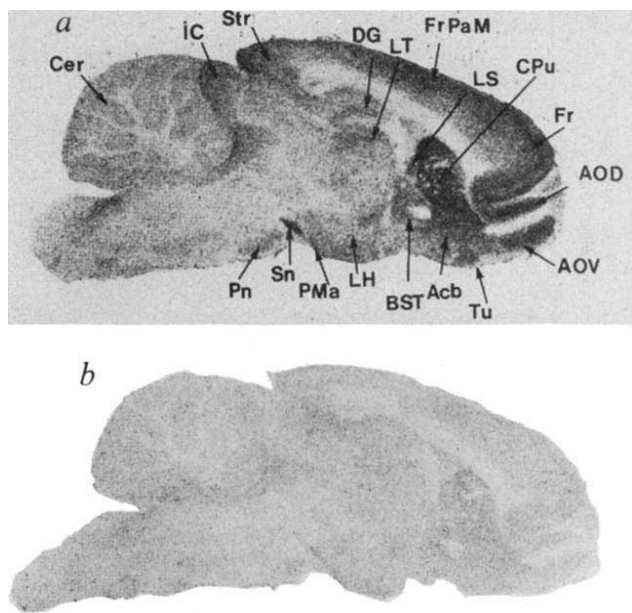


Fig. 3 Autoradiographic visualization of H_3 -receptors in rat brain using [3H](R) α -MeHA. Sagittal sections performed at level L: 1.4 mm of the stereotaxic atlas of Paxinos and Watson⁵⁰. *a*, Total binding in the presence of 1 nM [3H](R) α -MeHA; *b*, non-specific binding obtained by addition of 1 μ M thioperamide in the incubation medium. Note the low and rather uniform distribution of autoradiographic grains in *b*. Similar pictures were obtained when thioperamide was replaced by 30 nM (R) α -MeHA whereas 30 nM (S) α -MeHA did not significantly modify the [3H]ligand binding (not shown). Abbreviations: Acb, nucleus accumbens; AOD, anterior olfactory nucleus dorsal; AOV, anterior olfactory nucleus ventral; BST, bed nucleus of stria terminalis; Cer, cerebellum; CPu, caudate putamen; DG, dentate gyrus; Fr, frontal cortex; FrPaM, frontoparietal motor cortex; IC, inferior colliculus; LH, lateral hypothalamic area; LS, lateral septum; LT, lateral thalamus; PMa, perimammillary area; Pn, pontine nuclei; Sn, substantia nigra; Str, striate cortex; Tu, olfactory tubercle.

Methods. Male Wistar rats (180–200 g) were killed by decapitation, their brains rapidly removed, dropped into cold freon (-40°C monochlorodifluoromethane, Prestogaz R22) for a few minutes and kept frozen at -20°C until used. Sagittal frozen sections (20 μm thick) were prepared in a Ames cryostat at -18°C , thaw-mounted onto gelatin-coated glass slides and stored at -20°C until needed. Incubations (45 min at 25°C) were performed on slide-mounted tissue sections by dipping them into 1 nM [3H](R) α -MeHA in 50 mM phosphate buffer pH 7.4 containing 0.2% gelatin. Non-specific binding was obtained on immediately adjacent sections by adding 1 μM thioperamide in the incubation medium. The incubations were ended by dipping the slides briefly in drug-free ice-cold buffer and the tissue sections were rinsed sequentially in two 3-min phosphate buffer baths, then dipped into distilled water to remove any salts, immediately dried and apposed to [3H]Ultrafilms (LKB) in X-ray cassettes for 12 weeks at 4°C . Films were developed with Kodak D₁₉ (18 $^\circ\text{C}$; 5 min), fixed in Kodak Unifix (10 min) and rinsed under running tap water (30 min).

nucleus of the stria terminalis, and olfactory nuclei, to which diffuse projections of the ascending histaminergic neurons have been described³, showed the highest grain densities. In contrast the cerebellum (all layers), brainstem or mesencephalon (except a few areas, such as the substantia nigra), which contain a lower density of projections, showed fainter labelling. In the posterior hypothalamus a thin band of dense labelling was observed in the perimammillary area which is known to contain most histamine perikarya^{3,27,28} suggesting the presence of H_3 -receptors on the latter. However in the remainder of hypothalamus, in which levels of L-histidine decarboxylase and histamine are much higher than in telencephalon³, a relatively low degree of labelling

was observed. This may indicate that H_3 -receptors density is not the same on all histamine neurons and/or that H_3 -receptors are not restricted to the latter.

Conclusions and prospects

The potent and selective compounds that we describe here have already allowed us not only to confirm the existence of a third subclass of histamine receptors in brain but also to locate them and to elucidate their physiological involvement in the control of histamine synthesis and release in the living animal. In this respect the function of H_3 -receptors seems to be analogous to that of other classes of presynaptic receptors mediating feedback regulation of several neurotransmitters^{34–36}. Among these, α_2 -adrenoreceptors seem negatively coupled to adenylate cyclase³⁷ and it would be of interest to assess whether H_3 -receptors function similarly. The novel pharmacological tools, particularly the antagonist which appears to be the first agent able to elicit a marked facilitation of histaminergic transmissions in the central nervous system, should be useful, in behavioural and other studies, for defining further the functions of histamine systems. It has been suggested that the latter may be involved in the control of states of sleep and wakefulness^{38–40}, cerebral circulation⁴¹, energy metabolism⁴² and hypothalamic hormone secretion⁴³, but hitherto, verification of these hypotheses has been hindered largely by the lack of adequate pharmacological tools.

An application of the novel agents that is already important is the identification of H_3 -receptors outside the central nervous system in tissues such as the lung in which they probably control mast-cell histamine formation and, possibly, release. Should this be the case, it might be of great practical importance for the understanding and treatment of allergic and inflammatory mechanisms, to establish whether other mediators, coexisting with histamine in mast cells, are also controlled by H_3 -receptors.

Finally, using the novel agents, H_3 -receptors with a pharmacology closely similar to those characterized in laboratory animals have been unambiguously identified in human brain (J.-M.A., M.G., J.-P. Chodkiewicz and J.-C.S., manuscript in preparation). This suggests that these, or similar agents, may find therapeutic applications.

This work was supported in part by a grant from the Direction des Recherches Etudes et Techniques. We thank Mrs V. Chenal and D. Saveaux for their technical and secretarial help.

Received 15 December 1986; accepted 23 March 1987.

1. Plaut, M. & Lichtenstein, L. M. in *Pharmacology of Histamine Receptors* (eds Ganellin, C. R. & Parsons, M. E.) 392–435 (Wright PSG, Bristol 1982).
2. Hakanson, R. & Owman, C. *Life Sci.* **6**, 759–766 (1967).
3. Schwartz, J.-C., Garbarg, M. & Pollard, H. in *Handbook of Physiology, the Nervous System, Vol. IV: Intrinsic Regulatory Systems of the Brain* (eds Bloom, F. E., Mountcastle, W. B. & Geiger, S. R.) 257–316 (American Physiological Society, Bethesda, 1986).
4. Happola, O., Soinila, S., Paivarinta, H., Panula, P. & Eranko, O. *Histochemistry* **82**, 1–3 (1985).
5. Berridge, M. J. & Irvine, R. F. *Nature* **312**, 315–321 (1984).
6. Hegstrand, L. R., Kanof, P. D. & Greengard, P. *Nature* **260**, 163–165 (1976).
7. Black, J. W., Duncan, W. A. M., Durant, C. J., Ganellin, C. R. & Parsons, M. E. *Nature* **236**, 385–390 (1972).
8. Arrang, J.-M., Garbarg, M. & Schwartz, J.-C. *Nature* **302**, 832–837 (1983).
9. Arrang, J.-M., Garbarg, M. & Schwartz, J.-C. *Neuroscience* **15**, 553–562 (1985).
10. Arrang, J.-M., Garbarg, M. & Schwartz, J.-C. in *Frontiers in Histamine Research* (eds Ganellin, C. R. & Schwartz, J.-C.) 143–153 (Pergamon, Oxford, 1985).
11. Ganellin, C. R. in *Pharmacology of Histamine Receptors* (ed Ganellin, C. R. & Parsons, M. E.) 10–102 (Wright PSG, Bristol, 1982).
12. Schunack, W., Schwarz, S., Gerhard, G., Bücktimkin, S. & Elz, S. in *Frontiers in Histamine Research* (eds Ganellin, C. R. & Schwartz, J.-C.) 39–46 (Pergamon, Oxford, 1985).
13. Arrang, J.-M., Schwartz, J.-C. & Schunack, W. *Eur. J. Pharmacol.* **117**, 109–114 (1985).
14. Schwartz, J.-C., Arrang, J.-M., Garbarg, M. & Korner, M. *J. exp. Biol.* **124**, 203–224 (1986).
15. Arrang, J.-M. *et al. Eur. J. Pharmacol.* **111**, 73–84 (1985).
16. Gaddum, J. H. *J. Physiol., Lond.* **89**, 7P–9P (1937).
17. Cheng, Y. C. & Prusoff, W. H. *Biochem. Pharmacol.* **22**, 3099–3108 (1973).
18. Baudry, M., Martres, M. P. & Schwartz, J.-C. *Brain Res.* **116**, 111–124 (1976).
19. U'Prichard, D. C., Greenberg, D. A. & Snyder, S. H. *Molec. Pharmacol.* **13**, 454–473 (1977).
20. Yamamura, H. I. & Snyder, S. H. *Molec. Pharmacol.* **10**, 861–867 (1974).
21. Pert, C. B. & Snyder, S. H. *Science* **179**, 1011–1014 (1973).
22. Martres, M. P., Bouthenet, M. L., Sales, N., Sokoloff, P. & Schwartz, J.-C. *Science* **228**, 752–755 (1985).
23. Peroutka, S. J. & Snyder, S. H. *Molec. Pharmacol.* **16**, 687–699 (1979).
24. Lyesen, J. E., Nimegeers, C. J. E., Tollenaere, J. P. & Laduron, P. M. *Nature* **272**, 168–171 (1978).

25. Schwartz, J.-C., Arrang, J.-M., Garbarg, M. & Schunack, W. in *Innovative Approaches in Drug Research* (ed. Harms, A. F.) 73-89 (Elsevier, Amsterdam, 1986).
26. Garbarg, M., Barbin, G., Roderigas, E. & Schwartz, J.-C. *J. Neurochem.* **35**, 1045-1052 (1980).
27. Watanabe, T. *et al. Neurosci. Lett.* **39**, 249-254 (1983).
28. Steinbusch, H. W. M. & Mulder, A. H. in *Frontiers in Histamine Research* (eds Ganellin, C. R. & Schwartz, J.-C.) 119-130 (Pergamon, Oxford, 1985).
29. Eyre, P. & Chand, N. in *Pharmacology of Histamine Receptors* (eds Ganellin, C. R. & Parsons, M. E.) 298-322 (Wright PSG, Bristol, 1982).
30. Pearce, F. L. *et al. in Frontiers in Histamine Research* (eds Ganellin, C. R. & Schwartz, J.-C.) 411-422 (Pergamon, Oxford, 1985).
31. Hill, S. J., Emson, P. C. & Young, J. M. *J. Neurochem.* **31**, 997-1004 (1978).
32. Tran, V. T., Chang, R. S. L. & Snyder, S. H. *Proc. natn. Acad. Sci. U.S.A.* **75**, 6290-6294 (1978).
33. Gajtkowski, G. A., Norris, D. B., Rising, T. J. & Wood, T. P. *Nature* **304**, 65-67 (1983).
34. Langer, S. Z. *Br. J. Pharmacol.* **60**, 481-497 (1977).
35. Starke, K. A. *Rev. Pharmac. Tox.* **21**, 7-30 (1981).
36. Stjärne, L. in *Chemical Neurotransmission 75 Years. Karolinska Institutet, Nobel Conference No. 2* (eds Stjärne, L., Hedqvist, P., Lagercrantz, H. & Wennmalm, A.) 257-272 (Academic, London, 1981).
37. Schoffelemeier, A. N. M., Wierenga, E. A. & Mulder, A. H. *J. Neurochem.* **46**, 1711-1717 (1986).
38. Schwartz, J.-C. *A. Rev. Pharmac. Tox.* **17**, 325-339 (1977).
39. Vanni-Mercier, G., Sakai, K. & Jouvett, M. *C. r. hebd. Séanc. Acad. Sci., Paris* **298**, 195-200 (1984).
40. Lin, J.-S., Sakai, K. & Jouvett, M. *C. r. hebd. Séanc. Acad. Sci., Paris* **303**, 469-474 (1986).
41. Gross, P. M. in *Frontiers in Histamine Research* (eds Ganellin, C. R. & Schwartz, J.-C.) 341-352 (Pergamon, Oxford, 1985).
42. Quach, T. T., Duchemin, A. M., Rose, C. & Schwartz, J.-C. *Molec. Pharmacol.* **17**, 301-308 (1980).
43. Tuomisto, J. & Mannisto, P. *Pharmac. Rev.* **37**, 249-332 (1985).
44. Gerhard, G. & Schunack, W. *Arch. Pharm., Weinheim* **313**, 709-714 (1980).
45. Schunack, W. *Arzneimittel-Forsch.* **28**, 2199-2202 (1978).
46. Kollonitsch, J. *et al. Nature* **274**, 906-908 (1978).
47. Garbarg, M., Pollard, H., Quach, T. T. & Schwartz, J.-C. in *Methods in Biogenic Amines Research* (eds Parvez, S., Nagatsu, T., Nagatsu, I. & Parvez, H.) 623-662 (Elsevier, Amsterdam, 1983).
48. Sacki, K., Oishi, R. & Nishibori, M. in *Frontiers in Histamine Research* (eds Ganellin, C. R. & Schwartz, J.-C.) 173-184 (Pergamon, Oxford, 1985).
49. Parker, R. B. & Waud, D. R. *J. Pharmac. exp. Therap.* **177**, 1-12 (1971).
50. Paxinos, G. & Watson, C. *The Rat Brain in Stereotaxic Coordinates* (Academic, London, 1982).

LETTERS TO NATURE

Gravitational-wave bursts with memory and experimental prospects

Vladimir B. Braginsky* & Kip S. Thorne†

* Physics Faculty, Moscow State University, Moscow, USSR

† Theoretical Astrophysics, California Institute of Technology, Pasadena, California 91125, USA

Experimenters usually divide the gravitational waves which they hope to detect into three classes: 'bursts' in which the wave field h_{ij}^{TT} rises from zero, oscillates for only a few cycles and then returns to zero; 'periodic waves', and 'stochastic waves'¹. There is, however, a fourth class, 'bursts with memory'²⁻⁶ (BWM), in which h_{ij}^{TT} rises from zero, oscillates for a few cycles, and then after a burst of duration Δt settles down into a non-zero final value δh_{ij}^{TT} . Here we show that for any kind of detector the best way to search for a BWM is to integrate up the signal for an integration time $\hat{\tau} \approx 1/f_{opt}$, where f_{opt} is the frequency at which the detector has optimal amplitude sensitivity to ordinary bursts (bursts without memory). In such a search the sensitivity to BWM with duration $\Delta t \leq 1/f_{opt}$ is independent of the burst duration Δt and is approximately equal to the sensitivity to ordinary bursts one cycle long with frequency f_{opt} (see Fig. 1). It is possible, though not highly probable, that BWM will be among the earliest kinds of gravitational waves detected; therefore experimenters should take them into account when planning their search strategies and data analyses.

The reason that the optimum integration time is $\hat{\tau} \approx 1/f_{opt}$ is simple: if a BWM barely rises above the noise in the data after some integration time $\hat{\tau}$, then one cannot know what its actual duration Δt was; one can only know that there was some net change δh_{ij}^{TT} in the wave field during the time $\hat{\tau}$. This means that, independently of Δt (for $\Delta t < \hat{\tau}$), the BWM will appear like a BWM of duration $\hat{\tau}$, which in turn will be only marginally distinguishable from an ordinary burst of one cycle length and duration $\hat{\tau}$. The optimal sensitivity to such an ordinary burst occurs, by definition, at the frequency f_{opt} . Thus, the optimal sensitivity to a BWM must be for $\hat{\tau} \approx 1/f_{opt}$ and must equal the ordinary burst sensitivity at frequency f_{opt} .

This argument is independent of the type of detector used: resonant bar, laser interferometer on Earth or in space, monitoring of the Earth's seismic motions, skyhook, Doppler tracking of spacecraft, or timing of pulsars. However, for each type of detector there may be factors of order two to be gained, in searches for BWM, by careful planning of data analysis (optimal signal processing) and by careful planning of the experimental setup.

Figure 1 shows the sensitivities (for unity signal-to-noise ratio) of some present gravity-wave detectors and the limiting sensitivities of some detectors that have been proposed for future

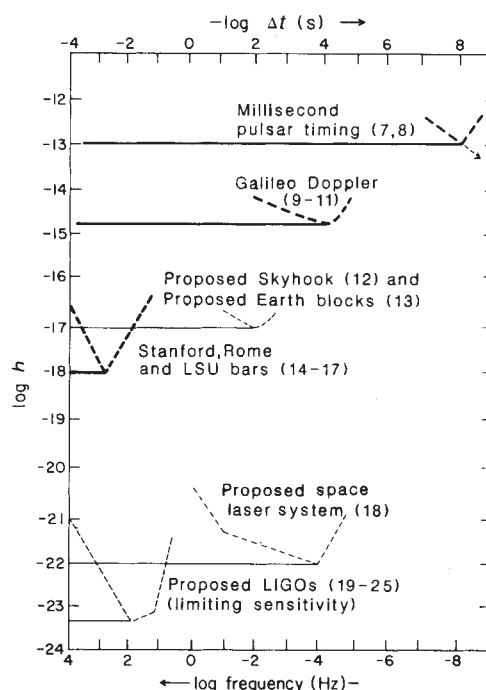


Fig. 1 Sensitivities of present (thick curves) and proposed (thin curves) gravitational-wave detectors. Dashed curves are sensitivities (dimensionless gravity-wave amplitude h) to ordinary bursts of one cycle duration, $\Delta t = 1/(\text{frequency})$. Solid curves are sensitivities to bursts with memory. Reference numbers are given in parentheses.

construction. The dashed curves are sensitivities to ordinary bursts; the solid curves, sensitivities (according to the above rule-of-thumb) to BWM. The thick curves are detectors now operating or due to operate in the near future. The thin curves are detectors proposed for construction and operation in the 1990s or (for laser system in space) in $\sim 2,000$. By 'sensitivity to ordinary bursts' (dashed curves) we mean sensitivity to a sinusoidal signal at the given frequency (bottom axis), which turns on abruptly, lasts for one cycle, then turns off abruptly. Details of the operating or proposed detectors will be found in the following references: search for fluctuations in the period of the millisecond pulsar relative to atomic clocks (primary frequency standards) (now underway)^{7,8}, Doppler tracking of the Galileo spacecraft (due to launch by NASA in late 1980s)⁹⁻¹¹, proposed Earth-orbiting skyhook¹², proposed multi-station monitoring of seismic vibrations of blocks of the Earth's crust with sizes 50 to 70 km¹³, Stanford, Rome, and Louisiana State University cryogenically cooled, resonant aluminium bars (now operating)¹⁴⁻¹⁷, proposed laser interferometer detector in space with 100 mW of laser power and with drag-free satellites at

corner and ends of 10^6 km arms¹⁸, and proposed Earth-based laser interferometer gravitational wave observatories (LIGOs) for which the curve shown corresponds to the limiting sensitivity with 100 W of laser power, 4-km arms, mirror reflectivities 0.9999, and seismic isolation down to 10 Hz¹⁹⁻²⁶.

It is clear that the lowest-frequency detectors can span the greatest range of BWM pulse durations Δt and for this reason they are of particular interest. (The poorer h -sensitivities of low-frequency detectors are compensated, at least in part, by the fact that the expected h -signal strengths are greater at lower frequencies^{1,26}.)

The prospects for low-frequency detectors are reasonably good. There is considerable room for improvement of micro-wave-frequency Doppler tracking of spacecraft, beyond the Galileo sensitivity; but improvements will depend on development of improved secondary frequency standards (present frequency stability $\Delta\omega/\omega \approx 3 \times 10^{-16}$ for 300 s averaging time^{27,28}; short-term goals both for cooled hydrogen masers²⁸⁻³⁰ and for superconducting cavity oscillators³¹ $\Delta\omega/\omega \approx 1 \times 10^{-17}$; ultimate quantum limit³² $\Delta\omega/\omega \approx 3 \times 10^{-20} (\hat{\tau}/300 \text{ s})^{-1/2}$). Improvements in Doppler tracking will also require multi-frequency tracking to remove charged-particle dispersion, and either putting the tracking antenna in space or closely monitoring neutral-particle dispersion in the Earth's troposphere^{33,34}. Laser-interferometer tracking of drag-free satellites by each other¹⁸ looks very promising in theory but is many years in the future. Pulsar timing^{7,8} extends to the longest Δt of all techniques. The limiting value, $(\Delta t)_{\max} = \hat{\tau} = 1/f_{\text{opt}}$ is set by the finite time for which the pulsar's period has been monitored, and recently the h -sensitivity has reached the limits of stability of primary atomic frequency standards. Thus, future improvements in $(\Delta t)_{\max}$ and h -sensitivity will require both continued monitoring of the pulsar to span a longer total time base, and improvements in primary frequency standards^{7,8}.

The different detectors respond to a BWM in fundamentally different ways, and this gives rise to different constraints on their performances. Resonant detectors (bars, skyhook, and Earth blocks) respond to a BWM that has duration $\Delta t \leq 0.5/(\text{resonant frequency})$ with an immediate displacement which thereafter remains constant until, after $\sim 1/4$ of the resonant period, it triggers long-lasting oscillations. However, if $\Delta t \gg 0.5/(\text{resonant frequency})$, the memory part of the burst has no significant effect.

Doppler tracking of spacecraft and pulsar timing respond to the 'gravitational redshift' aspect of a gravitational wave. Electromagnetic signals are used to compare the frequencies of clocks (pulsar rotation; atomic clocks on Earth or spacecraft). There is a redshift if and only if the gravitational-wave field is different at the emitting clock and receiving clock at the times of emission and reception, that is, if and only if the BWM sweeps through the electromagnetic wave during its propagation³⁵. This means that the BWM signal in the detection system lasts no longer than the light travel time between emitter and receiver. This accounts for the cutoff in sensitivity in Doppler tracking of Galileo (Fig. 1) at $\Delta t = (\text{round trip time to satellite}) \sim 10^4$ s. For pulsar timing the travel time is so long that it is irrelevant for human experimenters.

Laser interferometers respond to the proper-distance separations between their (ideally) free mirrors; they register the integral of h along the path between the mirrors. Therefore, if their mirrors were truly free, they would experience a truly permanent deformation when a BWM passes; in principle they can store the signal forever.

We turn now to a brief discussion of sources of BWM. In order to produce a BWM, an astrophysical system, either before the emission of the burst or afterwards or both, must consist of two or more freely moving masses²⁻⁵. These masses could be, for example, stars, black holes, gravitationally bound clusters of stars or black holes, or bursts of neutrinos; the only constraint is that each of the masses must move freely and therefore be

characterized by a single, constant 4-momentum. Thus, the source must consist either of the collision of two or more initially free masses, or an explosion of an initial single mass into several freely and independently moving masses. In either case, so long as the source is not at a cosmologically large distance, the permanent change in the gravitational-wave field (the burst's memory) $\delta h_{ij}^{\text{TT}}$ is equal to the 'transverse, traceless (TT) part'³⁶ of the time-independent, Coulomb-type, $1/r$ field of the final system minus that of the initial system. If $\mathbf{P}^A$ is the 4-momentum of mass A of the system and P_i^A is a spatial component of that 4-momentum in the rest frame of the distant observer, and if $\mathbf{k}$ is the past-directed null 4-vector from observer to source, then $\delta h_{ij}^{\text{TT}}$ has the following form:

$$\delta h_{ij}^{\text{TT}} = \delta \left(\sum_A \frac{4P_i^A P_j^A}{\mathbf{k} \cdot \mathbf{P}^A} \right)^{\text{TT}} \quad (1)$$

Here we use units with $G = c = 1$. In the observer's local Cartesian coordinate system oriented so the source is in the z -direction at the moment of emission, and in terms of the distance r to the source as measured in the observer's frame at the moment of emission and the properties of body A (mass M_A , components v_A^j of ordinary velocity, and energy per unit mass $\gamma_A = (1 - v_A^2/c^2)^{-1/2}$), and in ordinary c.g.s. units, this translates into

$$\delta h_{xx}^{\text{TT}} = -\delta h_{yy}^{\text{TT}} = \sum_A \frac{2GM_A}{c^4 r} \delta \left[\frac{\gamma_A (v_A^x v_A^x - v_A^y v_A^y)}{1 + v_A^2/c^2} \right] \quad (2a)$$

$$\delta h_{xy}^{\text{TT}} = \delta h_{yx}^{\text{TT}} = \sum_A \frac{4GM_A}{c^4 r} \delta \left[\frac{\gamma_A v_A^x v_A^y}{1 + v_A^2/c^2} \right] \quad (2b)$$

all other components vanish.

Correspondingly, for typical sources which presumably have γ_A not much larger than unity, the magnitude of the components of $\delta h_{ij}^{\text{TT}}$ can be estimated from the equation

$$\delta h \sim \frac{4G}{c^4} \frac{E_{\text{CM}}^{\text{kin}}}{r} \sim \frac{2Gm/c^2}{r} \frac{v^2}{c^2} \quad (3)$$

Here $E_{\text{CM}}^{\text{kin}}$ is the kinetic energy of the free masses in the system's centre of mass reference frame, m is the mass of the system's second heaviest free mass and v is its speed of motion relative to the heaviest mass. Numerically this gives for a source in our galaxy ($r \approx 10$ kpc) $\delta h \sim 10^{-17} (v/c)^2 (m/M_\odot)$; and for an object in the Virgo cluster of galaxies ($r \approx 10$ Mpc) $\delta h \sim 10^{-20} (v/c)^2 \times (m/M_\odot)$; and for an object at the Hubble distance (edge of observable universe, $r \approx 3$ Gpc) $\delta h \sim 10^{-20} (v/c)^2 (m/300 M_\odot)$. The burst duration is the time during which the gravity of the largest mass, M , significantly affects the velocity of the second largest mass, $\Delta t \sim GM/v^3 \sim (10^{-5} \text{ s}) (M/M_\odot) (v/c)^{-3}$.

Two promising examples of sources are: (i) a supernova explosion in which the stellar envelope and/or a $\sim$ one-second-long burst of neutrinos carries off the neutron-star binding energy, $E_{\text{CM}}^{\text{kin}} \approx 0.1 M_\odot c^2$, in a highly asymmetric manner³⁷. (What fraction of supernovae have highly asymmetric ejections is not known, but it may well be small.) Such an explosion would produce $\delta h \sim 10^{-18}$ if it were in our Galaxy and $\delta h \sim 10^{-21}$ if it were in the Virgo cluster; the burst duration might be $\Delta t \sim 1$ s. (ii) The interaction of three massive black holes in the nucleus of a galaxy at the Hubble distance, resulting in the ejection of one black hole of mass m at very high velocity v . Such an ejection would produce $\delta h \sim 10^{-20} (m/3000 M_\odot) (v/0.3 c)^2$ with $\Delta t \sim 1 \text{ s} (M/3,000 M_\odot) (v/0.3 c)^{-3}$. Such events might occur as part of the evolution which produces a supermassive black hole in a galactic nucleus^{38,39}. Just one such event per galaxy larger than the Magellanic clouds per age of the universe would lead to an event rate of roughly one per year.

These and other examples, together with the limiting sensitivities of proposed detectors as shown in Fig. 1, suggest that BWM could be among the first sources detected.

This work was supported in part by the Ministry of Higher Education of the USSR and by the NSF of the USA.

Received 25 August 1986; accepted 26 February 1987.

1. Smarr, L. (ed.) *Sources of Gravitational Radiation* 7-35; 477-497 (Cambridge University Press, 1979).
2. Zel'dovich, Ya. B. & Polnarev, A. G. *Soviet Astr.* **18**, 17-23 (1974).
3. Smarr, L. *Phys. Rev. D* **15**, 2069-2077 (1977).
4. Bontz, R. J. & Price, R. H. *Astrophys. J.* **228**, 560-575 (1979).
5. Kovacs, S. J. & Thorne, K. S. *Astrophys. J.* **224**, 62-85 (1978).
6. Braginsky, V. B. & Grishchuk, L. P. *Soviet Phys. JETP* **62**, 427-430 (1985).
7. Davis, M. M., Taylor, J. H., Weisberg, J. M. & Backer, D. C. *Nature* **315**, 547-550 (1985).
8. Taylor, J. H. *Proc. 11th International Conf. General Relativity and Gravitation Stockholm*, July 1986 (ed. MacCallum, M.) (Cambridge University Press, in the press).
9. Estabrook, F. B. *Acta Astronaut.* (in the press).
10. Bertotti, B. & Carr, B. J. *Astrophys. J.* **236**, 1000-1011 (1980).
11. Hellings, R. *Phys. Rev. D* **23**, 832-843 (1981).
12. Braginsky, V. B. & Thorne, K. S. *Nature* **316**, 610-612 (1985).
13. Braginsky, V. B., Gusev, A. V., Mitrofanov, V. P., Rudenko, V. N. & Yakimov, V. N. *Usp. fiz. Nauk* **147**, 422-424 (1985).
14. Fairbank, W. M. et al. in *Abstr. Contributed Papers Eleventh Int. Conf. on General Relativity and Gravitation Vol. II*, 580 (University of Stockholm, 1986).
15. Amaldi, E. et al. in *Abstr. Contributed Papers Eleventh Int. Conf. on General Relativity and Gravitation Vol. II*, 592 (University of Stockholm, 1986).
16. Hamilton, W. O. et al. in *Abstr. Contributed Papers Eleventh Int. Conf. on General Relativity and Gravitation Vol. II*, 582 (University of Stockholm, 1986).
17. Michelson, P. in *Gravitational Radiation* (ed. Deruelle, N. & Piran, T.) 465-474 (North-Holland, Amsterdam, 1983).
18. Faller, J. E., Bender, P. L., Hall, J. L., Hils, D. & Vincent, M. A. *Proc. Colloq. on Kilometric Optical Arrays in Space Cargese, Corsica*, 23-25 October 1984, 157-163 (ESA publication SP-226, April 1985).
19. Drever, R. & Weiss, R. in *Abstr. Contributed Papers Eleventh Int. Conf. on General Relativity and Gravitation Vol. II*, 605 (University of Stockholm, 1986).
20. Hough, J. et al. in *Abstr. Contributed Papers Eleventh Int. Conf. on General Relativity and Gravitation Vol. II*, 605 (University of Stockholm, 1986).
21. Maischberger, K. et al. in *Abstr. Contributed Papers Eleventh Int. Conf. on General Relativity and Gravitation Vol. II*, 598 (University of Stockholm, 1986).
22. Newton, G. P. et al. in *Proc. of Fourth Marcel Grossman Meet. Rome*, June 1985 (ed. Ruffini, R.) (North-Holland, Amsterdam, in the press).
23. Spero, R. et al. in *Proc. Fourth Marcel Grossman Meet. Rome*, June 1985 (ed. Ruffini, R.) (North-Holland, Amsterdam, in the press).
24. Shoemaker, P. et al. in *Proc. Fourth Marcel Grossman Meet. Rome*, June 1985 (ed. Ruffini, R.) (North-Holland, Amsterdam, in the press).
25. Winkler, W. et al. in *Proc. Fourth Marcel Grossman Meet. Rome*, June 1985 (ed. Ruffini, R.) (North-Holland, Amsterdam, in the press).
26. Thorne, K. S. in *Three Hundred Years of Gravitation* (ed. Hawking, S. W. & Israel, W.) (Cambridge University Press, in the press).
27. Stein, S. R. & Turneaure, J. P. in *Future Trends in Superconductive Electronics, Proc. AIP Conf. No. 44* (ed. Deaver, B. S.) 192 (American Institute of Physics, New York, 1978).
28. Smarr, L. L., Vessot, R. F. C., Lundquist, C. A. & Decker, R. *Gen. Relativ. Grav.* **15**, 129-163 (1983).
29. Walsworth, R. L. et al. *Phys. Rev. A* **34**, 2550-2552 (1986).
30. Vessot, R. F. C. in *Radio Astronomy in Space, Proc. Greenbank Workshop October 1986* (National Radio Astronomy Observatory, Green Bank, West Virginia, in the press).
31. Dick, G. J. & Strayer, D. M. *Proc. 38th Annual Frequency Control Symp.* 435-446 (USAERADCOM, Fort Monmouth, New Jersey, 1984).
32. Braginsky, V. B. & Viatchanin, S. P. *Soviet Phys. JETP* **47**, 433-435 (1978).
33. Armstrong, J. W. & Sramek, R. A. *Radio Sci.* **17**, 1579-1586 (1982).
34. Resch, G. M., Hogg, D. E. & Napier, P. J. *Radio Sci.* **19**, 411-418 (1984).
35. Estabrook, F. B. & Wahlquist, H. D. *Gen. Relativ. Grav.* **6**, 439-447 (1975).
36. Misner, C. W., Thorne, K. S. & Wheeler, J. A. *Gravitation* (Freeman, San Francisco, 1973).
37. Nadyozhin, D. K. in *Sources of Gravitational Radiation* (ed. Smarr, L.) 477-479 (Cambridge University Press, 1979).
38. Rees, M. J. in *Gravitational Radiation* (eds Deruelle, N. & Piran, T.) 297-320 (North-Holland, Amsterdam, 1983).
39. Saslaw, W. C., Valtonen, M. & Aarseth, S. J. *Astrophys. J.* **190**, 253-270 (1974).

Surface density of faint high-redshift quasi-stellar objects

Scott F. Anderson* & Bruce Margon†

* Mount Wilson and Las Campanas Observatories, 813 Santa Barbara Street, Pasadena, California 91101-1292, USA

† Department of Astronomy, FM-20, University of Washington, Seattle, Washington 98195, USA

The majority of known very high-redshift ($z \geq 3.4$) quasi-stellar objects (QSOs) have relatively bright apparent magnitudes^{1,2}, $R \leq 18$, and recent systematic surveys for fainter high-redshift QSOs³ have failed to find such objects, perhaps implying that the epoch of QSO formation is luminosity dependent. The $z = 4.01$ QSO recently reported⁴ is quite faint, and was selected by colour rather than spectrum, raising the possibility that previous slitless spectroscopic surveys have been biased against high-redshift QSOs. Here we report the discovery of a faint, red QSO ($B \geq 21$, $R \sim 19$) at $z = 3.56$, selected by 'traditional' low dispersion spectroscopic methods. The surface density of very faint, high-redshift QSOs may not be negligible, with a variety of lines of evidence suggesting that there are at least $\sim 0.2 \text{ deg}^{-2}$ objects with $z \geq 3.4$.

Although the number counts of QSOs as a function of apparent magnitude are now reasonably well determined to faint limiting magnitudes, this is certainly not the case for the redshift distribution. It has been recognized for many years that very high-redshift QSOs ($z > 3.4$) are rare, and that the small number of known cases have remarkably bright apparent magnitudes, typically $R < 18$. This effect has been highlighted by the recent successes of Hazard and collaborators² in discovering a substantial fraction of all such objects currently known by use of UK Schmidt plates of modest limiting magnitude. Whether the brightness of most known high z objects is due to simple observational selection, or indicates something physically interesting about the evolution of QSOs, has until recently been highly uncertain, although there was a suspicion that at least some fainter high- z QSOs would have been found if they exist⁵. This suspicion has been quantified by recent work of Schmidt *et al.*³, who have reported the results of deep, carefully calibrated QSO surveys over small regions of sky. They conclude that there is indeed a severe absence of faint high- z objects compared with extrapolations of existing QSO luminosity functions and that this implies that the high- z cutoff of the QSO distribution is

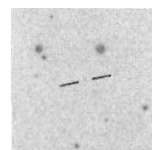


Fig. 1 A finding chart for the QSO 1409+732, enlarged from a portion of the red (E) print of the Palomar Observatory Sky Survey (© 1960 National Geographic Society-Palomar Sky Survey; reproduced by permission of the California Institute of Technology); the field shown is a 4 arc min square, with north up and east to the left. Although we do not expect it to be prominent in this reproduction, the object, located between the indicator bars, is clearly present on the red Survey print and absent on the blue one.

sharply luminosity dependent, presumably a fundamental boundary condition for theories of QSO evolution.

During the course of a project unrelated to this issue⁶, we have surveyed a relatively large area of sky, 20 deg^2 , to faint magnitudes, using low dispersion slitless spectra at the 4-m telescopes of the Kitt Peak National Observatory and Cerro Tololo Interamerican Observatory, and the Canada-France-Hawaii Telescope. The observations, which employed a 'grism' or 'grens' as the dispersing element, fell into three groups: (1) the entire 20 deg^2 was surveyed with the III-aF photographic emulsion and blue sensitive dispersers; (2) roughly 3 deg^2 of the same area were surveyed with the IV-N emulsion and red-sensitive dispersers, following the approach of Koo and Kron⁷; and (3) about 3 deg^2 of area were surveyed with the III-aF and red grism combination, similar to the approach of Osmer⁵. This last series of plates omitted the filter which restricted Osmer's results⁵ to $\lambda \geq 5,700 \text{ Å}$. In the 3 deg^2 of sky surveyed in which both red and blue grism data on III-aF plates are available, we are sensitive to Lyman- α QSOs (objects for which Ly- α is redshifted into the appropriate passband) throughout the redshift range $z = 1.7-4.7$, to continuum magnitudes of order 20, and observed emission line equivalent widths of $\geq 200 \text{ Å}$. This limiting equivalent width, of course, depends, for example, on the continuum magnitude and seeing⁸. The QSOs of our sample in the redshift range $z = 1.8-3.0$ have been discussed elsewhere^{6,9}, but recently we have begun follow-up spectroscopy of the $z > 3$ candidates. Our limited observations so far have concentrated on candidates selected from the red grism and III-aF combination, and from such a plate, obtained on Kitt Peak on February 1, 1984, we report here the discovery of a

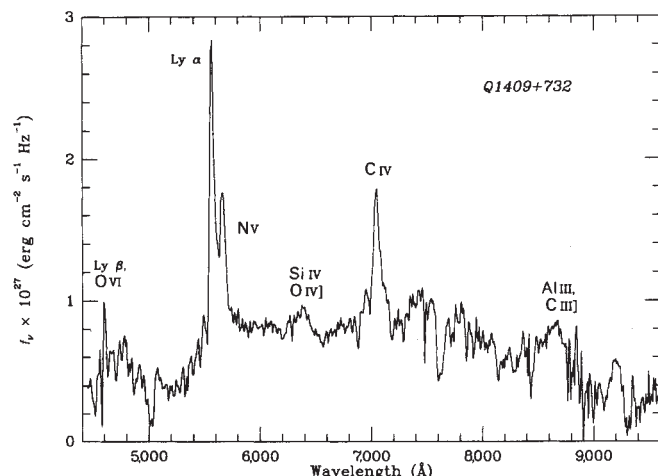


Fig. 2 The spectrum of the QSO 1409+732, obtained in a 30-min exposure with the Mayall 4-m reflector of Kitt Peak, using the Cryogenic Camera. Prominent emission features are identified. The telluric A band is responsible for the absorption near $\lambda 7600$.

faint, previously uncatalogued QSO, 1409+732, with a redshift of $z = 3.56$.

Figure 1 shows a finding chart for 1409+732, reproduced from the Palomar Observatory Sky Survey red print. The coordinates of the object are $\alpha(1950) = 14^{\text{h}} 09^{\text{m}} 11.9^{\text{s}}$, $\delta = 73^{\circ} 13' 34''$, with an uncertainty of $\pm 10''$. Although seen on the Sky Survey red print, the object is virtually undetected on the corresponding blue print. We roughly estimate $B \geq 21$ from the absence of an image there. The faintness of the object in the blue is also reflected on our grism plate material. On our III-aF blue grism plate of the field, the spectrum is almost non-existent, although the plate is quite deep; however, on the red grism/III-aF plate, the spectrum is well exposed, and emission lines corresponding in wavelength to Ly- α /N V and C IV at $z \sim 3.6$ are strongly detected.

On June 4 1986, we obtained a slit spectrum of 1409+732, using the 4-m telescope of the Kitt Peak National Observatory, with the Cryogenic Camera, a grism spectrograph with charge-coupled device detector. The grism employed (no. 810) has 150 lines mm^{-1} , and yields wavelength coverage of 4,700–9,300 Å with 30 Å spectral resolution. The data were reduced to absolute flux units using observations of four standard stars, which were spread in time throughout the night. The resulting flux-calibrated spectrophotometry is displayed in Fig. 2. From this spectrum we derive a redshift of $z = 3.56 \pm 0.01$ for 1409+732, with emission at the Ly- α /N V complex, C IV $\lambda 1549$, and Si IV $\lambda 1400$ clearly evident. The low sensitivity of the detector in the blue allows us to say little about the spectral characteristics of 1409+732 at wavelengths shortward of Ly- α in the QSO rest frame. However, there appears to be evidence of absorption by the Ly- α 'forest' clouds, or perhaps a damped Ly- α feature; the faintness of the object on the Sky Survey blue print also suggests these possibilities, and/or the presence of a Lyman limit continuum discontinuity. Although noisy, the data further indicate the presence of the Ly- β /O VI emission blend; these features are apparently detected on our red grism plate as well.

Comparison of the flux calibrations obtained when applying the various standard stars separately suggests an uncertainty in the absolute calibration of our data of ~ 0.5 magnitude. From our spectrophotometry we derive for the QSO a continuum magnitude of $m(1,475 \text{ Å}) = 19.1$ at an observed wavelength $\lambda_{\text{obs}} = 1,475 \text{ Å}(1+z)$, in the magnitude system described by Osmer¹⁰. Using relations given in ref. 11, and the assumption of a power law continuum in the optical with slope $\alpha_{\text{opt}} = 0.5$, we estimate that for 1409+732, the absolute B magnitude is $M_B = -28.0$ for $q_0 = 0.1$ and $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

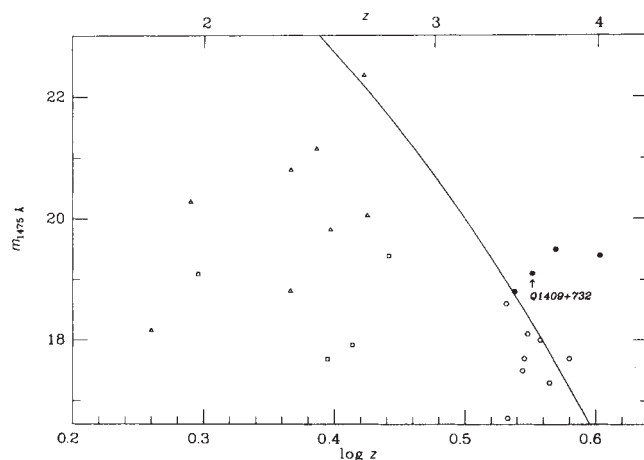


Fig. 3 The continuum magnitude at $1,475 \text{ Å}$, $m_{1,475 \text{ Å}}$, against the logarithm of the redshift for various QSO samples. Δ , $\square$, QSOs from the two surveys of Schmidt *et al.*³ (a few objects of $z < 1.5$ have been omitted for convenience); solid line, functional relationship, consistent¹⁴ with the surveys of Schmidt *et al.*³, to describe the evolution of z_{max} with luminosity; $\circ$, QSOs of $z \geq 3.4$ from the literature; $\bullet$, the four particularly faint high redshift QSOs noted in the text, including the new object reported here.

There are several fainter stellar objects located ≤ 30 arcs south-west of the QSO, and on our grism plates, their poorly detected spectra appear not dissimilar to that of the QSO. Motivated by the interesting possibility that this system might be a very high redshift gravitational lens, we obtained the spectrum of the brightest of these nearby images, which is also closest to the QSO, offset 13 arcs to the south-west. The spectrum shows the absorption features of a normal galactic star.

As an aside, we note that 1409+732 lies within the 'Ursa Minor Deep Survey' area, a region which has been imaged to very sensitive X-ray flux levels with the Einstein Observatory. This QSO is not detected in the X-ray images (F. Primini, personal communication), with a 4.5σ X-ray flux upper limit of $4.5 \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$ (0.5–3 keV), which translates into a lower limit on the X-ray to optical flux ratio¹² of $\alpha_{\text{ox}} > 1.52$. This X-ray limit is not particularly remarkable, especially given the apparent decrease in the ratio of X-ray to optical luminosity for high optical luminosity and/or high redshift QSOs¹³.

For redshifts as high as that of 1409+732, most previous data imply that QSOs of comparably faint optical luminosity are exceedingly rare. Schmidt *et al.*³ interpret their null detections of faint high redshift QSOs as indicative of a luminosity-dependent cutoff in the QSO redshift distribution. In Fig. 3, we have plotted redshift against apparent magnitude for 1409+732, the QSOs from the two Schmidt *et al.* surveys, and a functional form for the redshift cutoff¹⁴ (assuming $q_0 = 0.1$ and $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$) that is consistent with the Schmidt *et al.* data. For convenient comparison, we have converted Gunn r magnitudes for the Schmidt *et al.* QSOs into $m(1,475 \text{ Å})$ magnitudes, under the assumption that $\alpha_{\text{opt}} = 0.5$. We also include in Fig. 3 analogous data, where available in the literature, for all other QSOs with $z \geq 3.4$. The Schmidt *et al.* data and the associated model suggest that the redshift cutoff for $M_B = -28.0$ is at $z_{\text{max}} \leq 3.4$; alternatively, for these data and models, a QSO of redshift $z \sim 3.6$ should be brighter than $m(1,475 \text{ Å}) = 18.3$. From Fig. 3, it is clear that 1409+732 is, by comparison, unusually faint.

Of the approximately fourteen catalogued QSOs with $z > 3.4$, many have $m < 17$ –18, but there are a few others which are comparably faint to 1409+732. For the radio-selected PKS1351–018¹⁵ and grism-selected 2227–394^{10,16}, we estimate $m(1,475 \text{ Å}) = 19.6$ ($M_B = -27.6$) and $m(1,475 \text{ Å}) = 18.8$ ($M_B = -28.2$) respectively. Warren *et al.*⁴ report yet another example

of a very faint ($B \sim 21$) colour-selected QSO, 0046–293, with $z = 4.01$, for which we adopt $m(1,475 \text{ \AA}) = 19.4$. In Fig. 3 we see that, from the viewpoint of some current data and models, these objects (although of $M_B = -27$ to -28), like 1409+732, are also quite faint. It is somewhat surprising, given the observational selection effects which make identification of such very faint objects difficult, that $\sim 25\%$ of the currently known QSOs with $z > 3.4$ are, in fact, so faint, despite the previously mentioned beliefs that all very high redshift objects are abnormally luminous. The three most discordant objects in Fig. 3, including the object newly reported here, have been found by three different selection techniques, and this suggests to us that the faintness of these QSOs, rather than a particular selection bias, may largely account for the difficulty in identifying them. Figure 3 also implies that current characterizations^{3,14} of the functional form of the faint, high redshift end of the QSO Hubble diagram may be in need of revision.

As noted above, the observed surface density of faint, very high redshift objects is quite uncertain. Our survey for these QSOs is just beginning, but we have identified at least one such object in the 3 deg^2 surveyed with both red and blue grism III-aF plates. Similarly, the Hoag and Smith¹⁶ survey, using a blue grism and Kodak 127-04 (predecessor to the III-aF) plate combination, found one such QSO, 2227–394, in a search area of 5.1 deg^2 . Thus, these admittedly meagre data suggest that there may be (at least) of order 0.2 deg^{-2} QSOs of $z > 3.4$ and $m(1,475 \text{ \AA}) \leq 19.5$. Additionally, at least one such object, PKS1315–018 is present¹⁵ in a radio-selected sample which surveyed $\sim 85 \text{ deg}^2$; although extrapolation to such high redshifts is uncertain, if we conservatively adopt the value of < 0.1 for the ratio of radio-loud to radio-quiet QSOs¹⁷, the discovery of PKS1315–018 is also consistent with such a surface density.

Based on extrapolations from counts at lower redshifts, Hazard and McMahon¹ predicted a surface density of $\sim 0.2 \text{ deg}^{-2}$ for QSOs in the redshift range $3.3 < z < 3.9$ and $R < 19.3$, in good agreement with the rough estimate given here. Further, that our surface density estimate for objects with $z > 3.4$ and $m(1,475 \text{ \AA}) \leq 19.5$ is similar to the $\sim 0.1 \text{ deg}^{-2}$ which Hazard *et al.*² find for $3.3 < z < 3.9$ QSOs of $R \leq 18$, lends support to their suggestion that the luminosity function for such high redshift QSOs is flat at these magnitudes. Finally, we note that such a low surface density is consistent with the null detection of faint, very high-redshift QSOs in modest area covered by the Schmidt *et al.*³ surveys; nor are the null detections of such faint QSOs in the redshift ranges $z = 3.7\text{--}4.7$ and $z = 3.9\text{--}6.2$ in the areas surveyed by, respectively, Osmer⁵ and Koo and Kron⁷, in conflict with such surface densities.

Numerous selection effects may apply to the type of QSO discussed here. Dunlop *et al.*¹⁵ have noted the extreme red colour of PKS1315–018, $(B - R) = 1.6$, and 1409+732 seems likely to have similar colours (at least in $B - R$). As Dunlop *et al.* point out, such colours may inhibit easy selection of such very high redshift QSOs in the traditional colour-selected surveys. Further, because of their very red colours, objects such as 1409+732 and PKS1315–018 could be expected to appear only at very faint magnitudes in any QSO samples whose limiting completeness magnitudes are defined in terms of broadband blue magnitudes (for example, B or J). For example, 1409+732 and PKS1315–018 would not appear in such a sample complete to $B = 20.5$, despite their measured values of $R \sim 19$. PKS1315–018 has rather weak emission lines¹⁵; for example, the rest frame equivalent width of Ly- α in PKS1315–018 is only $35 \text{ \AA}$, whereas the mean value for the mainly radio-selected QSOs studied by Wilkes¹⁸ is $65 \text{ \AA}$ (with N v contributing, on average, an additional $19 \text{ \AA}$ rest frame equivalent width when blended). Similarly, we find that 1409+732 also has intrinsically weak lines, with rest frame equivalent widths for C IV and Ly- α /N v, of 20 and $40 \text{ \AA}$, respectively. However, note that at such high redshifts the observed equivalent widths of Ly- α /N v and C IV

in both 1409+732 and PKS1315–018 are still quite substantial (see Fig. 2) because they scale as $(1+z)$ times the rest frame equivalent width. It certainly was these prominent emission lines that allowed easy grism selection of 1409+732. It has been suggested^{15,19} that a systematic decrease in emission line equivalent width with increasing redshift will seriously bias prism techniques against discovery of high redshift objects. In our view, whether or not the decrease of Ly- α /N v emission equivalent width is in fact faster than the $(1+z)$ increase factor remains to be proven.

Contrary to some earlier studies, we suggest that the technique of selecting faint, very high redshift QSOs using a red sensitive grism/detector combination does seem to hold substantial promise for mapping the cutoff in the redshift distribution, even at magnitudes $m(1,475 \text{ \AA}) \geq 19$, provided that a large area of the sky (of the order of 10 deg^2) is surveyed. The lack of success reported earlier with this approach may merely have been due to the circumstance that the areas surveyed were comparable in size to that in which one might expect to find ≤ 1 faint, high redshift QSO. Even if large-area surveys, such as the one we have undertaken, prove to be incomplete for the selection of these objects, they can at least provide interesting lower limits on the surface densities of such QSOs, whereas small area surveys with null detections can provide only approximate upper limits.

The authors were Visiting Astronomers at Kitt Peak National Observatory and Cerro Tololo Interamerican Observatory, National Optical Astronomy Observatories, operated by AURA Inc., under contract with the NSF; and at the Canada-France-Hawaii Telescope, operated by the NRC, CNRS and UH.

Note added in proof: M. Schmidt *et al.*, using techniques similar to ref. 3, report the discovery of two faint QSOs at $z > 3.4$ in a survey of 14 deg^2 (*Astrophys. J. Lett.*, in the press). The implied surface density agrees with that which we discuss.

Received 20 October 1986; accepted 17 February 1987.

1. Hazard, C. & McMahon, R. *Nature* **314**, 238–240 (1985).
2. Hazard, C., McMahon, R. G. & Sargent, W. L. W. *Nature* **322**, 38–40 (1986).
3. Schmidt, M., Schneider, D. P. & Gunn, J. E. *Astrophys. J.* **306**, 411–427 (1986); **310**, 518–533 (1986).
4. Warren, S. J. *et al.* *Nature* **325**, 131–133 (1987).
5. Osmer, P. *Astrophys. J.* **253**, 28–37 (1982).
6. Anderson, S. F. & Margon, B. *Astrophys. J.* **314**, 111–128 (1987).
7. Koo, D. C. & Kron, R. G. *Publ. astr. Soc. Pac.* **92**, 537–545 (1980).
8. Clowes, R. G. *Mon. Not. R. astr. Soc.* **197**, 731–738 (1981).
9. Anderson, S. F. thesis, Univ. Washington, Seattle (1985).
10. Osmer, P. S. *Astrophys. J. Suppl. Ser.* **42**, 523–540 (1980).
11. Schmidt, M. & Green, R. F. *Astrophys. J.* **269**, 352–374 (1983).
12. Tananbaum, H. *et al.* *Astrophys. J.* **234**, L9–L13 (1979).
13. Avni, Y. & Tananbaum, H. *Astrophys. J.* **262**, L17–L21 (1982).
14. Bechtold, J., Weymann, R. J., Lin, Z. & Malkan, M. A. *Astrophys. J.* **315**, 180–197 (1987).
15. Dunlop, J. S. *et al.* *Nature* **319**, 564–567 (1986).
16. Hoag, A. A. & Smith, M. G. *Astrophys. J.* **217**, 362–381 (1977).
17. Sramek, R. A. & Weedman, D. W. *Astrophys. J.* **221**, 468–480 (1978).
18. Wilkes, B. J. *Mon. Not. R. astr. Soc.* **218**, 331–361 (1986).
19. Trimble, V. & Woltjer, L. *Science* **234**, 155–161 (1986).

IRAS constraints on the sizes of Pluto and Charon

Edward F. Tedesco*, Glenn J. Veeder Jr*,
R. Scott Dunbar* & Larry A. Lebofsky†

* Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, USA

† Lunar and Planetary Laboratory, The University of Arizona, Tucson, Arizona 85721, USA

Recent mutual eclipse events have begun to yield information on the sizes of Pluto and Charon. Their detection by the Infrared Astronomical Satellite (IRAS) during its sky survey has provided an important additional constraint. Here we report how models of their thermal emission indicate that Charon contributes a significant amount of the observed infrared radiation. The most

probable diameters for Pluto and Charon are 2,200 and 1,300 (± 150) km. These results are consistent with there being some atmosphere on Pluto.

Since its discovery in 1930 Pluto has been an enigmatic object. The fact that it is considerably smaller than the other outer planets was immediately obvious because it failed to show a discernible disk even when observed through the largest telescopes of the time. Using its predisccovery predicted mass of 6.6 Earth masses together with an upper limit on its diameter set by its stellar appearance led to ridiculously high values for its density¹. Although as time passed the values derived for its mass continued to decrease it was not until the discovery of its satellite, Charon, in 1978 that a reliable mass was obtained and the realization made that even the most recent mass estimates^{2,3} had been too large by a factor of 50^{4,5}.

Up until the mid-1970s the best value for Pluto's diameter came from visual observations made by G. Kuiper in 1950 using the Palomar 200-inch telescope (5,900 km; ref. 6), and from a near-miss stellar occultation which set an upper limit of 6,800 km (ref. 7). The next downward revision in the estimated diameter came with the detection of methane frost on the surface of Pluto⁸. Diameter estimates for Pluto between 2,800 and 3,300 km were based on the assumption of an average geometric albedo between 0.4 and 0.6, by analogy with icy satellites. Speckle observations made between 1978 and 1980^{9,10} gave non-limb-darkened values between 3,000 and 4,000 km for the diameter.

The technique of thermal radiometry at wavelengths of 10 and 20 μm , which was applied so successfully for determining the diameters of asteroids and planetary satellites, could only set an upper limit for Pluto's diameter because to date no firm detection has been made with the most sensitive of ground-based infrared equipment. Application of an asteroid thermal model using a 22.5- μm flux upper limit determined by Lebofsky, Rieke and Lebofsky in 1982¹¹ resulted in an upper limit of 2,600 km for the diameter of Pluto. This diameter included a correction for the presence of Charon (L.A.L., unpublished).

The passage of the orbital plane of Charon through the Sun during 1987 and 1988 results in a series of mutual eclipse events centred around those years. Shallow events begun in 1984, were first observed in January and February 1985¹², and will continue through 1990¹³. The best current estimates of the dimensions of Pluto and Charon come from modelling mutual event light curves observed in 1985 and 1986^{13,14}.

On two dates in July 1983 IRAS detected Pluto during its all-sky survey. The mean monochromatic flux density at 60 μm , the only band in which a detection was obtained, was 0.42 Jy with a formal uncertainty of 0.04 Jy (co-added 60- μm flux data).

We have used several combinations of asteroid thermal models to estimate the range of parameters which yield this infrared flux and remain consistent with the visual mutual eclipse event models. Morrison and Lebofsky¹⁵, and references therein, describe the non-rotating (standard) and rotating asteroid thermal models. The standard model applies to a surface in equilibrium with solar radiation and has been successfully used to interpret icy satellite infrared data. The rotating model applies to a surface with very high thermal inertia. In addition, we have also considered the simple case of an isothermal spherical grey body.

The adopted model parameters include an emissivity of 0.9, an infrared beaming factor of 0.9 and a semi-empirical phase integral of 0.8–0.9. The mean absolute visual magnitudes of Pluto and Charon are -0.61 and $+0.98$, respectively, assuming that the mean absolute magnitude of the Pluto–Charon system is -0.80 ^{14,16,17} and using the size/albedo ratios given by Dunbar and Tedesco¹³. If the size/albedo ratios determined by Reinsch and Pakull¹⁴ are used these would become -0.56 and $+1.10$. Because the uncertainties in the thermal model diameters are dominated by the uncertainty in the 60 μm flux, we adopted -0.6 and $+1.0$. A series of self-consistent infrared models were generated assuming these values for the absolute visual magnitudes of Pluto and Charon.

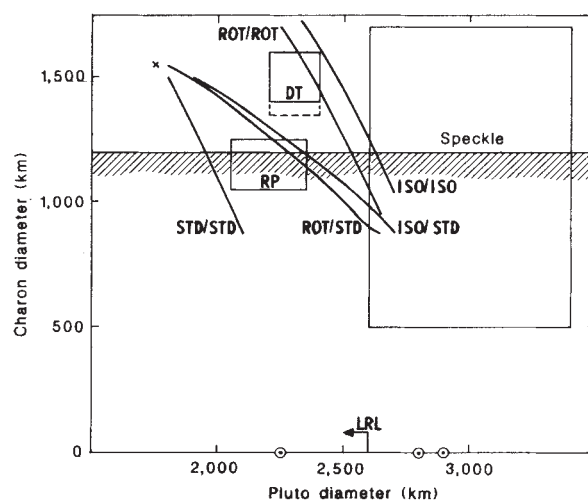


Fig. 1 A comparison of the diameters of Pluto and/or Charon based on various techniques. Speckle interferometry (box labelled speckle) gives the non-limb-darkened limits on the diameters of Pluto and Charon set by speckle observations by Arnold, Boksenberg and Sargent⁹, and by Hege *et al.*¹⁰; limb-darkened results yield larger diameters). Visual eclipse light curve modelling (boxes labelled DT¹³ and RP¹⁴, respectively). The dashed extension to the box labelled DT gives the results obtained applying the Dunbar and Tedesco¹³ model to the data set they used augmented by two additional observations published by Reinsch and Pakull¹⁴). Occultation of a star by Charon (cross-hatched horizontal line indicates the minimum diameter for Charon set by Walker's¹⁸ occultation observation). The ground-based thermal infrared measurements (line with arrow labelled LRL on abscissa) indicates the two-sigma upper limit to the diameter of Pluto set by Lebofsky, Rieke and Lebofsky's¹¹ failure to detect Pluto at 22.5 μm using the NASA Infrared Telescope Facility. The thermal radiometric diameters assuming all the 60- μm flux observed by IRAS originates from a standard model for Charon, that is, Pluto albedo of unity, (x in upper left corner) or all from either a standard, rotating or isothermal model for Pluto (circled dots on abscissa, respectively). Locus of diameters resulting from applying the standard (STD), rotating (ROT) and isothermal (ISO) models in the sense Pluto/Charon. For example, the line labelled ROT/STD indicates the diameter pairs permitted by the 60- μm flux using a rotating thermal model for Pluto and a standard (non-rotating) thermal model for Charon and, as in all cases, mean absolute visual magnitudes of -0.6 for Pluto and $+1.0$ for Charon.

tudes of Pluto and Charon. The resulting paired ranges of diameters are compared graphically in Fig. 1 to diameter results derived from speckle interferometry, mutual-event light curve models, a stellar occultation by Charon and ground-based radiometry.

No composite of standard thermal models for both Pluto and Charon is consistent with the mutual event or speckle results. Indeed, if both Pluto and Charon are assumed to conform to the standard model then an upper limit on Pluto's diameter of 1,950 km results, assuming that Charon's diameter is greater than 1,200 km as required by Walker's occultation observation¹⁸. Composite thermal models with both Pluto and Charon rotating or both isothermal are more consistent with both the mutual event and speckle results. Composite thermal models with a non-rotating Charon and either a rotating or isothermal Pluto are most consistent with both sets of mutual event results. The uncertainty in the flux measured by IRAS results is an uncertainty of approximately ± 150 km in the model curves and so the distinction between rotating and isothermal models is only marginally significant.

Together, the composite infrared model results presented here suggest a lower limit to the diameter of Pluto of 1,750 km in the extreme case that its visual geometric albedo is unity and all the infrared radiation is emitted by Charon. The upper limit for

the diameter of Pluto (2,600 km) set by Lebofsky, Rieke and Lebofsky¹¹ is consistent with the mutual event modelling but not the speckle results. If the albedo of Charon is lower than that of Pluto, as suggested by the mutual event modelling, all the thermal models require that Charon be warmer than Pluto. The diameter of Charon must be at least as large as its observed occultation chord (1,200 km) and therefore Charon may contribute more of the infrared flux observed by IRAS than Pluto. Therefore, the best present estimates for the diameters of Pluto and Charon, taking into account the eclipse light-curve diameters, Charon occultation lower limit, Pluto ground-based infrared upper limit, and IRAS survey 60 μm flux measurement, are $2,200 \pm 150$ km, and $1,300 \pm 150$ km, respectively.

The failure of the standard model for Pluto implies that its thermal properties differ significantly from those of asteroids and the icy satellites of the jovian planets all of which are well described by this thermal model. Either Pluto must show a significant amount of exposed bare (solid) ice or its surface temperature distribution is distorted by the presence of an atmosphere¹⁹. Icy satellites appear to have regoliths of low thermal inertia. Therefore, the relatively low infrared flux observed from the Pluto-Charon system supports the suggestion that Pluto does have a significant atmosphere. Such an atmosphere is not required for Charon.

We thank Tom Chester for his help in obtaining the IRAS 60- μm flux data and H. H. Aumann for fruitful discussions. Part of this research was conducted at the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA. L.A.L. was supported by a NASA grant to the University of Arizona.

Received 14 January; accepted 24 February 1987.

- Lowell, P. *Mem. Lowell Obs.* 1, no. 1 (1915).
- Ash, M. E., Shapiro, I. I. & Smith, W. B. *Science* 174, 551 (1971).
- Seidelmann, P. K., Klepczynski, W. J., Duncombe, R. L. & Jackson, E. S. *Astr. J.* 76, 488-492 (1971).
- Christy, J. W. & Harrington, R. S. *Astr. J.* 83, 1005-1008 (1978).
- Duncombe, R. L. & Seidelmann, P. K. *Icarus* 44, 12-18 (1980).
- Kuiper, G. P. *Publ. astr. Soc. Pacif.* 62, 133 (1950).
- Halliday, I., Hardie, R. H., Franz, O. G. & Priser, J. B. *Publ. astr. Soc. Pacif.* 78, 113-124 (1966).
- Cruikshank, D. P., Pilcher, C. B. & Morrison, D. *Science* 194, 835-837 (1976).
- Arnold, S. J., Boksenberg, A. & Sargent, W. L. W. *Astrophys. J.* 234, L159-L163 (1979).
- Hege, K. *et al. Icarus* 50, 72-81 (1982).
- Lebofsky, L. A., Rieke, G. H. & Lebofsky, M. J. *Bull. Am. astr. Soc.* 14, 766 (1982).
- Binzel, R. P., Tholen, D. J., Tedesco, E. F., Buratti, B. J. & Nelson, R. M. *Science* 228, 1193-1195 (1985).
- Dunbar, R. S. & Tedesco, E. F. *Astr. J.* 92, 1201-1210 (1986).
- Reinsch, K. & Pakull, M. W. *Astr. Astrophys.* (in the press).
- Morrison, D. & Lebofsky, L. A. in *Asteroids* (ed. Gehrels, T.) 184-205 (University of Arizona Press, Tucson, 1979).
- Tedesco, E. F. & Tholen, D. J. *Bull. Am. astr. Soc.* 12, 729 (1980).
- Binzel, R. P. & Mulholland, J. D. *Astr. J.* 89, 1759-1761 (1984).
- Walker, A. R. *Mon. Not. R. astr. Soc.* 192, 47 (1980).
- Trafton, L. & Stern, S. A. *Astrophys. J.* 267, 872-881 (1983).

Pesticides in rainwater in the northeastern United States

R. Peter Richards, Jack W. Kramer, David B. Baker & Kenneth A. Krieger

Water Quality Laboratory, Heidelberg College, 310 East Market Street, Tiffin, Ohio 44883, USA

The use of agricultural herbicides in the United States increased 280% between 1966 and 1981; insecticide use also increased, but only by a few per cent¹. Since 1981, pesticide use has fluctuated with crop acreage, and shows no clear increasing or decreasing trend². Compared to their predecessors, currently used herbicides are not very toxic to animals, and both herbicides and insecticides are less persistent and show less tendency to bio-accumulate, but are more soluble in water, and therefore more mobile in the environment. Previous studies of these compounds have reported their movement to the edge of the field by surface transport in rainfall runoff³, their presence in surface and ground water⁴⁻⁹, and their seasonal presence in drinking water supplies⁴. By contrast, only one previous study¹⁰ has reported any of these compounds in rainwater: atrazine was reported in Maryland in concentrations as high as $2.19 \mu\text{g l}^{-1}$. The work we report here indicates that many of these compounds are more commonly present in rainfall, at least regionally, than has previously been realized.

A pilot study of rainwater samples taken at our laboratory in north-central Ohio during 1984 revealed the presence of several herbicides and insecticides in readily detectable amounts. In spring and summer 1985, samples were obtained at our laboratory, and also at West Lafayette, Indiana; Parsons, West Virginia and Potsdam, New York. Samples were collected in stainless steel basins which were manually uncovered during rainfall events. Except where several sequential samples were taken during an event, all samples are integrated over the period of rainfall. Analyses of 19 herbicides and insecticides were performed by dual-column capillary gas chromatography, using nitrogen-phosphorus thermionic specific detectors. The use of two different columns (Durabond-5, Durabond-1) allowed simultaneous confirmation of concentrations for 14 of the 19 compounds. Spiked samples indicated recoveries (Table 1) generally around 70%. Detection limits based on measurements of

Table 1 Detection limits and ranges of linear response, based on analysis of dilution series of mixed standards, and mean per cent recoveries of spikes, for herbicides and insecticides detected in rainfall in 1985

Pesticide	Detection limit $\mu\text{g l}^{-1}$	Mean recovery (%)	Linear range $\mu\text{g l}^{-1}$
Herbicides			
Alachlor (Lasso)	0.1	64	>0.5
Atrazine (Aatrex)	0.05	69	>0.5
Butylate (Sutan)	0.05	70	>0.2
Cyanazine (Bladex)	0.25	79	>0.5
EPTC (Eradicane, Eptam)	0.05	66	ND
Linuron (Lorox, Linurex)	1.5	80	>5.0
Metolachlor (Dual)	0.25	67	>0.25
Metribuzin			
(Sencor, Lexone)	0.1	65	>2.5
Penoxalin (Prowl)	0.05	71	ND
Simazine (Princep)	0.25	74	>2.5
Insecticides			
Carbofuran (Furadan)	0.2	77	>0.5
Fonofos (Dyfonate)	0.05	57	>0.15
Terbufos (Counter)	0.1	54	ND

ND, not determined.

dilution series of standards are listed in Table 1. Concentrations have been adjusted to 100% recovery using the mean percent recovery for each compound, but no adjustment has been made for nonlinearity at low concentrations. Reported concentrations which fall below the linear region indicated in Table 1 are low estimates of the true concentrations. Further details of sample collection and preparation and of analytical methodology are presented elsewhere¹³.

The data gathered in 1985 represent analyses of 79 samples for 19 compounds which represent 90% (by weight) of the pesticide use in Ohio¹⁴. Eleven of these compounds were detected in the rainwater samples. The Ohio samples represent 23 of the 30 rainfall events exceeding 3 mm per day, and 75% of the rainfall volume, for the months April through August. The concentrations we have observed in samples from Ohio and Indiana (up to $6 \mu\text{g l}^{-1}$) are much higher than peak rainwater concentrations reported¹¹ for dichloro-diphenyl-dichloro-ethene (DDT) ($0.01 \mu\text{g l}^{-1}$) and dieldrin ($0.03 \mu\text{g l}^{-1}$), and many are higher than those reported¹² for toxaphene ($0.497 \mu\text{g l}^{-1}$).

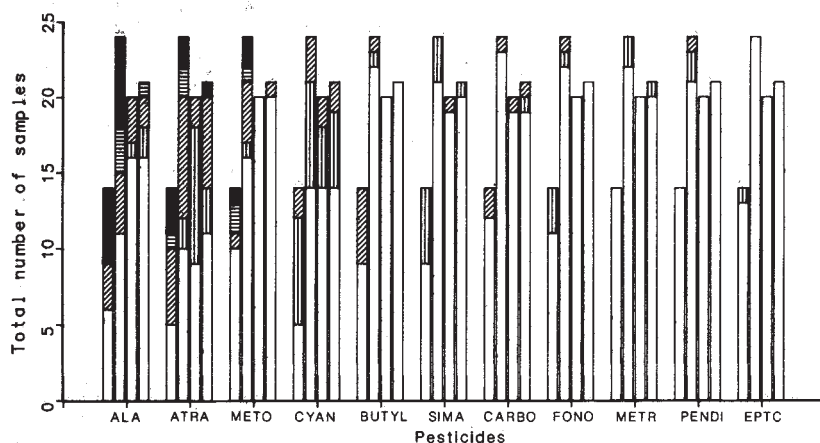


Fig. 1 Concentration ranges of herbicides and insecticides in rainfall in spring and summer 1985. From left to right within each set of four, the bars present data from West Lafayette, Indiana; Tiffin, Ohio; Parsons, West Virginia and Potsdam, New York. Unshaded, not detected; vertical shading, $<0.1 \mu\text{g l}^{-1}$; diagonal shading, $0.1-0.5 \mu\text{g l}^{-1}$; horizontal shading, $0.5-1.0 \mu\text{g l}^{-1}$; solid, $>1.0 \mu\text{g l}^{-1}$. Full pesticide names are listed in Table 1.

The data are presented in Figs 1 to 4, and suggest several preliminary conclusions.

Samples from the New York and West Virginia stations showed measurable concentrations of all compounds less frequently, and in lower concentration, than did the samples from Indiana and Ohio (Fig. 1). These stations are in areas of less intense pesticide use than the Indiana and Ohio stations, but they are downwind from the midwestern region, where pesticide use is generally high. The differences in concentration patterns suggest that local contributions to the atmospheric pool are important.

The presence of these compounds in rainfall is strongly seasonal (Fig. 2). Concentrations were below detection limit in early spring before application, reached maxima in the rainfall events following application in May, and declined to unquantifiable levels by the end of July. Occasional samples taken during the autumn and winter failed to reveal any pesticides. These findings are consistent with the timing of application (mid to late May) and with typical half-life data for the compounds (several weeks in most cases), but are in contrast with those of the Maryland study¹⁰, which reported comparable but highly variable atrazine concentrations in atmospheric deposition throughout the year following application in May.

Our findings cannot be explained as a consequence of changes

in rainfall. Rainfall amounts during the period of measurable pesticide concentrations show no time trend, and volume-adjusted deposition (rainfall multiplied by concentration) plots closely parallel those of concentration alone.

During periods of frequent or prolonged rainfall, concentrations decreased sharply (Fig. 3) as compared with the general rate of decrease (Fig. 2). Concentrations in the next rainfall generally were higher than those at the ends of these periods. This pattern is consistent with a scavenging effect of precipitation on the local atmospheric load of these compounds. The rebound of concentration in the next rainfall is consistent with the movement of new air masses into the area and/or renewal of the atmospheric pool by transport from the fields in the time between rainfall events.

When airborne pollutants have moderately stable concentrations over time, scavenging often leads to an inverse relationship between concentration and rainfall amount¹⁵. Such a relationship would be expected for pesticide concentrations as well, but could not be demonstrated by multiple linear regression techniques in the face of the strong (but nonlinear) seasonal decline in concentrations.

For the Ohio data there is a strong relationship between the amount of a particular pesticide used and the presence of the pesticide in rainfall (Fig. 4). For the 13 pesticides for which data are available, the amount of pesticide used is correlated with the maximum observed rainfall concentration, the frequency of concentrations above detection limit, the mean concentration of the pesticide, and the sum of rainfall $\times$ concentration for all storms (a total deposition estimate). These correlations are all significant at probabilities $P < 0.001$ by Pearson's correlation coefficient. Rank order statistics gave similar results.

These results are consistent with a conceptual model in which pesticides become available to the atmospheric pool following application in the spring, and remain available for about two months, in amounts which decrease over time as the pesticides break down. Replenishment of the atmospheric pool probably involves both volatilization and wind erosion of soil particles with adsorbed pesticides. Atmospheric entrainment during application is probably a minor factor, especially with pesticides which are incorporated into the soil. Loss from the atmospheric pool occurs by precipitation and by breakdown. The data are consistent with an atmospheric pool which is local to regional in scale. Some transport out of the source region by atmospheric movement must be assumed, however our data lack the spatial resolution to give any indication of the importance of this export relative to deposition essentially in the source area.

Our purpose in writing this report is to call attention to the frequent presence of herbicides and insecticides seasonally in rainfall, at least in Indiana and Ohio. Similar findings are to be expected in other areas of intense agriculture. The environmental significance of these findings is uncertain. Although acute animal toxicities are generally low, effects of long-term exposure are not well known. Threshold levels for direct effects of herbicides

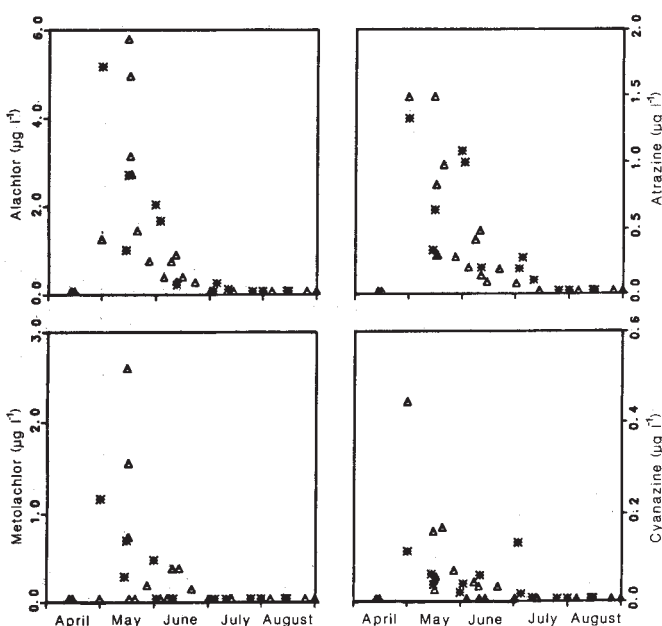


Fig. 2 Concentration of alachlor, atrazine, metolachlor and cyanazine in rain water, April through August 1985. Asterisks, samples from Indiana; triangles, samples from Ohio.

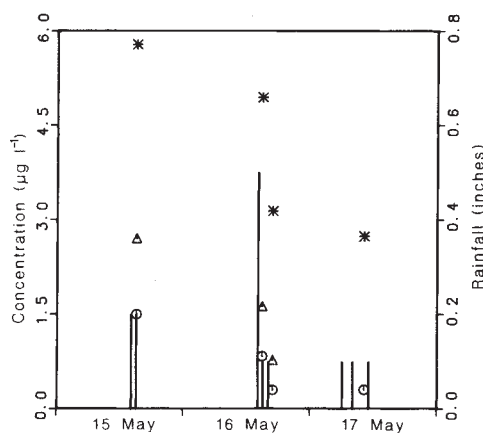


Fig. 3 Declining concentrations of alachlor (asterisks), metolachlor (triangles) and atrazine (circles) during a three-day rain period at Tiffin, Ohio, 15–17 May 1985. Hourly rainfall is shown by the vertical lines.

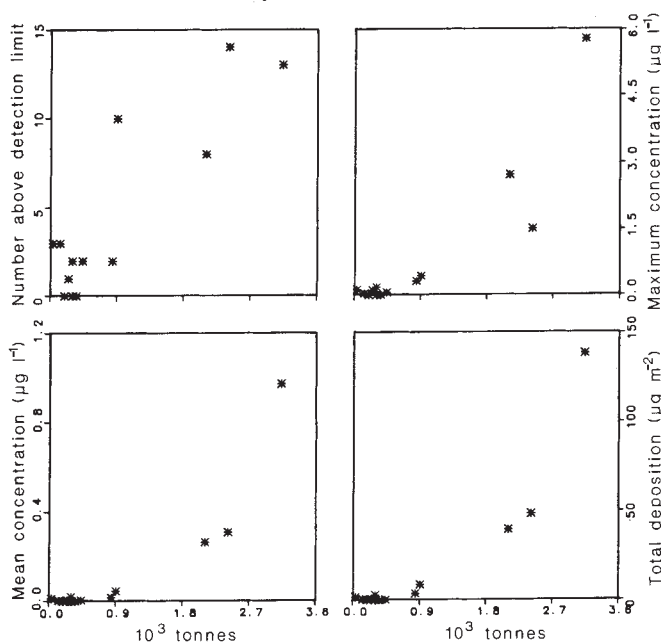


Fig. 4 Relationships between pesticide quantities used and measures of pesticide presence in rainfall. Values of r^2 range between 75% and 81%.

on most non-targeted plants are several orders of magnitude higher than concentrations observed in rainfall. Some aquatic algae and simple vascular plants show effects at or near these concentrations; however our data⁴ indicate that impacts on these species are much more likely from surface runoff than from rainfall. The atmospheric residence time of moderately stable pollutants¹⁵ indicates the possibility of widespread atmospheric dispersal of these compounds. Not enough is known about their pathways into and transport in the atmosphere, their temporal and spatial distribution patterns, and their partitioning in the atmosphere between vapor, water droplet solution, and sediment-adsorbed phases.

We thank Dr W. Moldenhauer, Dr David Helvey and Dr Thomas Young for collecting rainfall samples for us. This study was supported in part by the American Electric Power Service Corporation.

Received 16 September 1986; accepted 27 February 1987.

1. Hileman, B. *Envir. Sci. Technol.* **16**, 645A–650A (1982).
2. National Agricultural Chemicals Association Report to the Membership for 1984–1985 (Washington, 1985).
3. Wauchope, R. D. *J. envir. Qual.* **7**, 459–472 (1978).
4. Baker, D. B. *J. Soil Water Conserv.* **40**, 125–132 (1985).

5. Baker, D. B., Krieger, K. A., Richards, R. P. & Kramer, J. W. in *Perspectives on Nonpoint Source Pollution* (ed. Moore, M. L.) 210–207 (Environmental Protection Agency, Washington, 1985).
6. Frank, R., Braun, H. E., Holdrinet, M. V. H., Sirons, G. J. & Ripley, B. D. *J. envir. Qual.* **11**, 497–505 (1982).
7. Johnson, H. P. & Baker, J. L. *Field-to-stream Transport of Agricultural Chemicals and Sediment in an Iowa Watershed: Part II. Data Base for Model Testing (1979–1980)* (Environmental Protection Agency, Athens, Georgia, 1984).
8. Hallberg, G. R., Libra, R. D. & Hoyer, B. E. in *Perspectives on Nonpoint Source Pollution* 109–114 (Environmental Protection Agency, Washington, 1985).
9. Soren, J. & Stelz, W. G. in *Perspectives on Nonpoint Source Pollution* 101–108 (Environmental Protection Agency, Washington, 1985).
10. Wu, T. L. *Wat. Air Soil Pollut.* **15**, 173–184 (1981).
11. Tabatabai, M. A. in *Agricultural Management and Water Quality* (ed. Schaller, F.) 92–108 (Iowa State University Press, 1983).
12. Bidleman, T. F. & Christensen, E. J. *J. geophys. Res.* **84**, 7857–7862 (1979).
13. Kramer, J. W. & Baker, D. B. in *Quality Assurance for Environmental Measurements, ASTM STP 867* (eds Taylor, J. K. & Stanley, T. W.) 116–132 (American Society for Testing and Materials, Philadelphia, 1985).
14. Waldron, A. C. *Ohio agric. Res. Dev. Cent. Bull.* **1157**, 1–12 (Ohio State University, 1984).
15. Glass, G. E. & Brydges, T. C. in *Acid Rain/Fisheries* (ed. Johnson, R. E.) 265–286 (NE Division of American Fisheries Society, Bethesda, 1982).

Small synoptic/mesoscale eddies and energetic variability of the eastern Levantine basin

Allan R. Robinson*, Artur Hecht†, Nadia Pinardi*, Jonah Bishop†, Wayne G. Leslie*, Zvi Rosentroub†, Arthur J. Mariano* & Steve Brenner†

* Center for Earth and Planetary Physics, Harvard University, Cambridge, Massachusetts 02138, USA

† Israel Oceanographic and Limnological Research, Ltd, Haifa, Israel

Mesoscale eddies are energetically dominant over most of the world's ocean¹. The local dynamics of eddies is now a central problem of physical oceanography² and their statistics control aspects of the general circulation³. The horizontal scale of eddies is generally related to, but somewhat larger, than the Rossby internal deformation radius (the depth scale multiplied by the buoyancy frequency divided by the Coriolis frequency). In the eastern Mediterranean, which is not well explored or understood, the internal radius is known to be 12 km or 4 times smaller than in the North Atlantic. The study of such eddies requires very fine sampling which hinders their accidental discovery. By using a dedicated data set and objective analysis methods, we have established the existence of small synoptic/mesoscale eddies and jets, which constitute a variable and turbulent current regime and dominate the general circulation flow of the Levantine basin. Their existence in the Mediterranean has an impact on all aspects of basic and applied marine science there, both conceptually and with respect to experimental and monitoring programme designs.

The often cited general circulation data sets of Ovchinnikov and Fedorseyev⁴ and Ovchinnikov⁵ can reveal topographically controlled sub-basin-scale gyres, but do not hint at the mesoscale eddies. Even recent studies⁶ depict the currents of the eastern Levantine basin as a smooth broad continuation of the North African current, following the coasts of the basin. South-east of Cyprus the current divides into two branches: one continuing northwards into the Cilician basin; the other turning westwards towards the centre of the Levantine basin. Thus, classically the circulation of the eastern Levantine basin consists of a large cyclonic gyre between Cyprus and the Egyptian coast with a smooth flow of the order of 5–10 cm s⁻¹ between the surface and 500 m in all seasons.

A general circulation survey at 0.5° spacing in latitude and longitude⁷ provided the motivation for carrying out a dedicated search for the small mesoscale eddies. We obtained and analysed a fine resolution data set in the southeastern Levantine basin during October 1985. The ~200 km² region south of Cyprus is presented in Fig. 1a; the topography consists of a small south-north gradient with a seamount located south-west of Cyprus. We now believe this to be the 'extension region' of the North

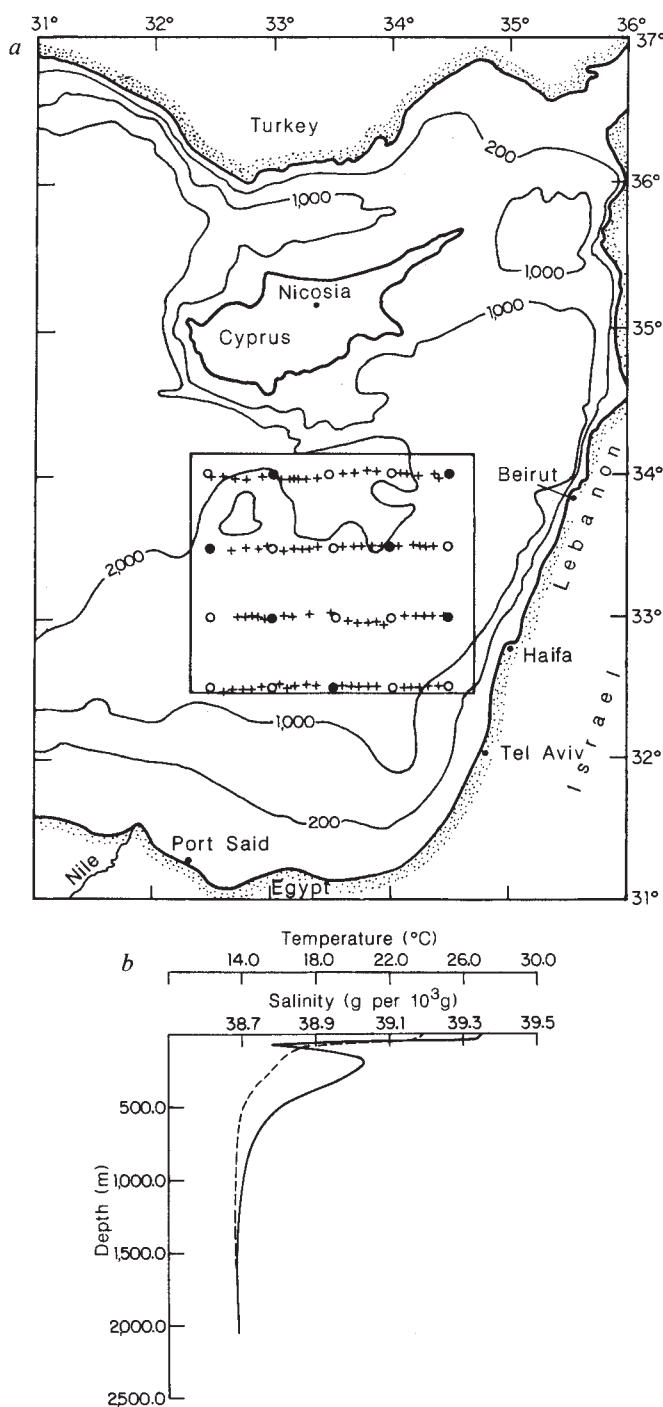


Fig. 1 *a*, The southeastern Levantine basin and bathymetric contours. The box indicates the domain of optimal interpolation. The circles show the CTD sampling scheme: open circles indicate casts down to about 1,000 dbar and full circles casts below 1,100 dbar. The crosses mark the positions of the XBT casts. *b*, The horizontally averaged temperature and salinity profiles for the combined XBT-CTD data set.

African current. The RV *Shikmona*, of the Israel Oceanographic and Limnological Research Institute, sampled this region with 82 expendable bathythermographs (XBTs) to 450 m at 10-km intervals (nominal) along the latitudinal tracks spaced 45 km apart and with twenty 1,100-m conductivity-temperature-depth (CTD) casts interspersed at 45-km intervals (every half degree of longitude) along the tracks. If a Rossby internal radius of deformation scaling were relevant, we would expect the diameter of the eddies to be about six times the internal radius. Thus, a sampling of the order of 10 km or less is expected to resolve

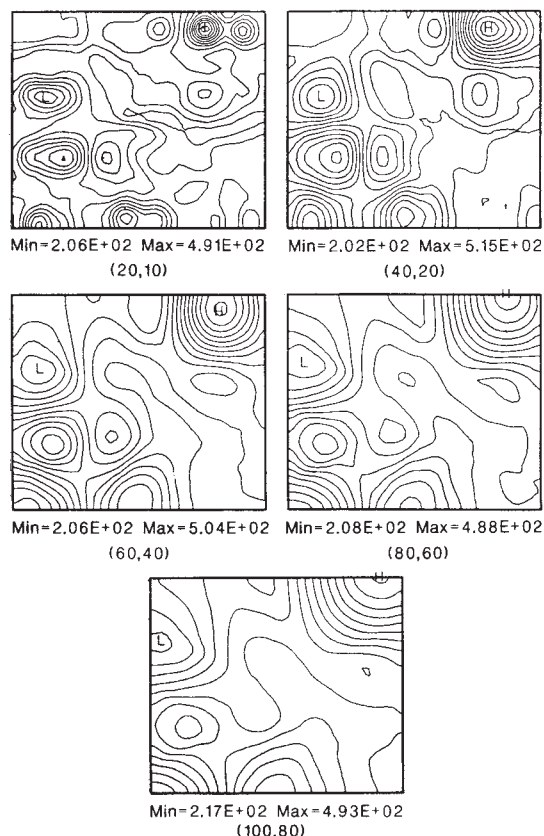


Fig. 2 Depth in metres of the 15°C isotherm for the combined XBT-CTD data set. The numbers below each map indicate the value of the correlation length (decay length in km) assumed in the objective mapping correlation function (see text). The contour interval is 20 m for each map. The numbers in parentheses denote the values of the parameters (*a*, *b*).

the eddies along the track. The requirement of high resolution sampling along tracks which are relatively coarsely separated, for the definition of mesoscale features, has been discussed elsewhere⁸.

We compared the XBT casts and the adjacent CTD temperature profiles; they agree within ± 10 dbar in the first 50 m of the water column and are practically identical between 50 and 450 m. These results are consistent with previous investigations elsewhere^{9,10}. The horizontally averaged temperature and salinity profiles from the combined CTD and XBT data sets are shown in Fig. 1*b*. The Levantine intermediate water layer, indicated by a subsurface salinity maximum, is found between 100 and 350 m with a relative salinity maximum at ~ 225 m. The temperature profile shows a very sharp temperature decrease in the first 100 m and a smaller main thermocline gradient down to 500 m. From comparison between repeated stations and examination of other data sets from the region, the high frequency internal wave noise on the temperature profiles is estimated to have an amplitude of the order of 5–10 dbar compared with the standard deviation of the mesoscale field presented here which is of the order of 50–60 dbar for the depth of the 15°C isotherm.

A series of objectively analysed maps was made from the high-resolution combined CTD and XBT data sets. The domain is $220 \times 180 \text{ km}^2$ and the grid spacing for the objective interpolation is 5 km. Since no previous mesoscale data sets exist, no statistics of the phenomena exist and therefore no *a priori* empirical correlation is available for objective analysis mapping¹¹. Therefore, we chose a simple homogeneous, isotropic two parameter model, $C(r) = (1 - r^2/a^2) \exp(0.5 r^2/b^2)$ with $r^2 = x^2 + y^2$ (ref. 12). To evaluate the parameters, (*a*, *b*), we

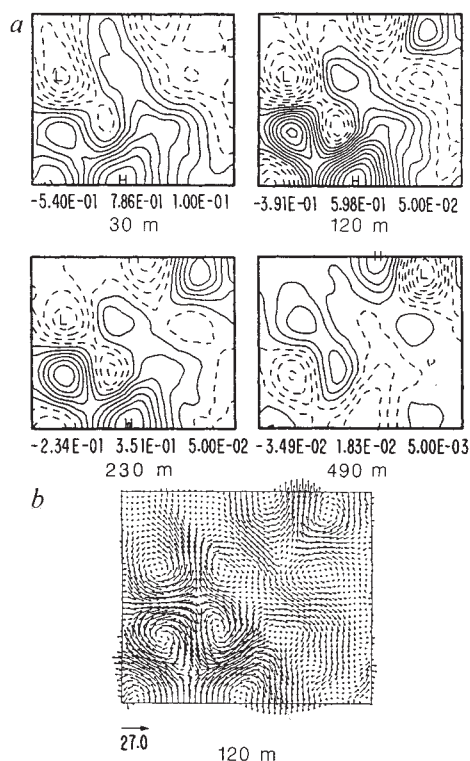


Fig. 3 *a*, Dynamic height anomaly at 30, 120, 230 and 490 m of depth: the horizontal average of the dynamic height at each level has been subtracted and the units are in $\text{m}^2 \text{s}^{-2}$. The values underneath each map represent the minimum, maximum and contour interval respectively. Solid contours indicate values above the horizontal average; broken contours indicate values below the horizontal average. The first solid contour equals the horizontal average. *b*, The geostrophic velocity structure at 120 m in units of cm s^{-1} . A reference arrow is shown at the left bottom corner.

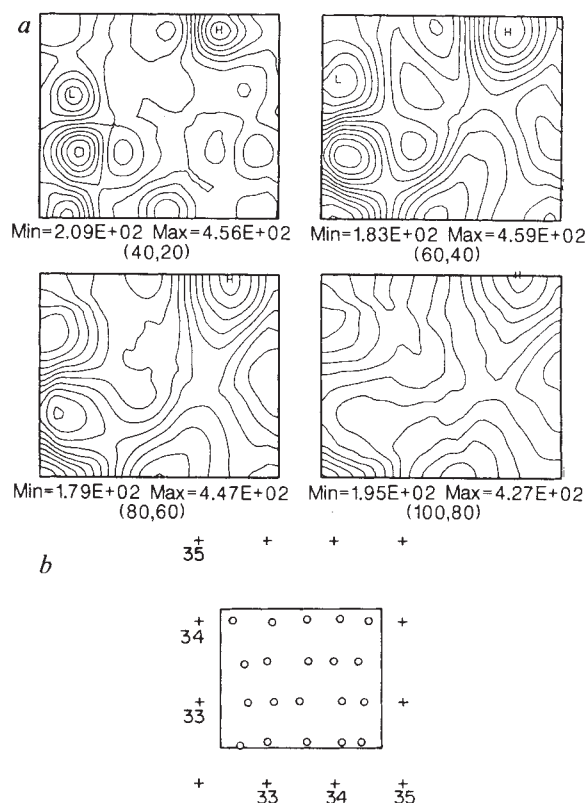


Fig. 4 *a*, Objective mapping of the depth of the 15°C isotherm for a subsampled XBT data set. The numbers in parentheses below the maps indicate the correlation function parameters (see text) used for the mapping. The contour interval is 20 m for each map. *b*, The subsampled data coverage: 20 measurements are left from a total of 82 fine sampling XBT measurements. The crosses indicate a latitude, longitude grid where the numbers represent North latitude and East longitude. The open circles are the positions of the observations.

objectively map the same data with a range of parameter values and compare these maps with each other and with those obtained by subjective mapping. The objective maps of the depth of the 15°C isotherm of Fig. 2 show that the most reasonable and consistent results are obtained with the choice of $(a, b) = (40, 20)$ or $(60, 40)$. If the correlation parameters are larger then the interpolated field is over-smoothed. If the parameters are smaller, the maps contain the same eddies (except for a split north-east corner eddy) but they are smaller, sharper and separated by quiescent regions, which seems less realistic. The difference between the $(40, 20)$ and $(60, 40)$ experiments is in the strength of the eddy field, but until a correlation function is deduced from the data themselves the two cases are equivalent. We shall choose to use $(60, 40)$ for further analysis and interpretation. For the parameters chosen each mapped eddy is defined by at least four observations.

The $(60, 40)$ field of Fig. 2 is populated by small eddies ~60–70 km in diameter and by a somewhat larger-scale eddy feature (not fully contained in the domain) in the north-east. The local internal Rossby radius of deformation, calculated from the first eigenvalue of the linearized vertical quasi-geostrophic operator¹³ using Fig. 1*b* profiles is 12.4 km. The small mesoscale field of Fig. 2 has a diameter 5–6 times the first internal Rossby deformation radius which is characteristic of such oceanographic phenomena generally. Maps of the temperature at several depths show the small eddies to have significant baroclinic amplitude at depths from 120 m to 330 m. Thus the small mesoscale eddy features have an intensified signature in the Levantine intermediate water layer.

In Fig. 3*a*, the dynamic height relative to 464 m is shown at several levels. This quite shallow reference level for the dynamic height has been chosen because it is the maximum depth com-

mon to all the XBT and CTD casts. The geopotential anomalies for XBT casts were computed from the XBT temperatures and the nearest CTD salinity. The dynamic height maps and the geostrophic velocity structure (Fig. 3*b*) show the intense small mesoscale eddies, embedded in a larger-scale flow composed of current filaments and jet segments. The North African current extension general circulation transport is probably contained in this flow but masked by the much stronger instantaneous mesoscale variability. The small mesoscale features have a strong vertically sheared structure with the baroclinic flow as intense as 20–30 cm s^{-1} in the upper thermocline levels. The near-surface 30-m geostrophic flow exhibits predominantly larger-scale structures.

The present data and analysis establish the existence of the eastern Mediterranean mesoscale but provide only one realization; the sampling requirements for future data sets are demanding. There does exist, however, in the region of Fig. 1*a* a several-year data set of >20 realizations obtained by the Israel Oceanographic and Limnological Research Institute scientists but with sampling essentially at the CTD locations only (the 'Marine Climate' or MC data). What, if any, information concerning mesoscale variability can be extracted from this data set? The sampling of 0.5° of latitude and longitude is regarded as 'fine general circulation sampling'; can it be regarded as 'coarse mesoscale resolution' sampling? To address this question we subsampled the October 1985 data set and intercompared maps of the subsampled data made with standard and coarse correlation functions. The results are shown in Fig. 4. We see that $(40, 20)$ and $(60, 40)$ maps do give a valid indication of the small mesoscale eddy field, although details differ from the best maps of Fig. 2 and only one or two data points define a feature.

This is, however, encouraging because of the possibility of extracting at least qualitative synoptic and statistical mesoscale information from the MC data set which is uniquely extensive for the eastern Mediterranean. Preliminary results indicate highly variable eddies, jets and filaments with mesoscale variability exceeding seasonal variability.

We have searched for and identified the existence of mesoscale eddy variability in the southeastern Levantine basin. These eddies are of an absolute size about 4 times smaller than, for example, eddies in the North Atlantic, but bear the same approximate scale ratio to the local Rossby internal deformation radius. Their dynamics and role in the general circulation and in transport is not known, nor are the characteristics of mesoscale eddies elsewhere in the eastern Mediterranean. First evidence for the existence of mesoscale eddies in the western Mediterranean has also been recently documented with timescales from 5 to 30 days by Taupier-Letage and Millot¹⁴; the space scale is not well defined but thought to be small. The ratio between the relative vorticity and planetary vorticity advection is ~ 10 for the scales of motion involved¹⁵; thus geostrophic turbulence effects are expected to be more important than planetary wave propagation. Moreover, we suspect that nonlinear interactions between mesoscale eddies with important higher vertical mode coupling, to be potentially important in the Mediterranean. In particular, the southeastern Levantine small mesoscale eddies are Levantine-intermediate-water-intensified and could have an important role in transporting those water properties throughout the basin. The kinematic picture which is emerging, of strong eddies, jets and filaments dominating the flow and masking underlying mean current transports is consistent with results from other oceanic regions, for example, the north-east Atlantic¹⁶ and the north-east Pacific⁸.

Our research and the measurements from October 1985 were made in the context of the POEM (physical oceanography of the eastern Mediterranean) Program, which is a cooperative investigation of the eastern Mediterranean involving general circulation and mesoscale studies of theoretical and experimental nature⁷ and which should provide a future context for determining the origin and role of the eddies. We thank the Intergovernmental Oceanographic Commission and Unesco Marine Science Program for helping to provide the POEM Scientific forums. We acknowledge the support of the Israel Oceanographic Limnological Research Institute and the NSF under a grant to Harvard University.

Received 27 October 1986; accepted 26 February 1987.

1. Robinson, A. R. in *Eddies in Marine Science* (ed. Robinson, A. R.) 1-14 (Springer, Heidelberg, 1983).
2. Haidvogel, D. B. in *Eddies in Marine Science* (ed. Robinson, A. R.) 404-436 (Springer, Heidelberg, 1983).
3. Holland, W. R., Harrison, D. E. & Semtner, A. J. Jr in *Eddies in Marine Science* (ed. Robinson, A. R.) 379-401 (Springer, Heidelberg, 1983).
4. Ovchinnikov, I. M. & Fedorseyev, A. F. in *Basic Features of the Geological Structure* 185-201 (1965).
5. Ovchinnikov, I. M. *Oceanology* 6, 49-59 (1966).
6. Lacombe, H. *News. coop. Invest. Mediter.* 7, 5-25 (1975).
7. POEM Steering Committee *UNESCO Reports in Marine Science* No. 35 (Unesco, Paris, France, 1985).
8. Moores, C. N. K. & Robinson, A. R. *Science* 223, 51-53 (1984).
9. Flierl, G. R. & Robinson, A. R. *J. phys. Ocean.* 7, 300-302 (1977).
10. Heinmiller, R. H. in *Eddies in Marine Science* (ed. Robinson, A. R.) 542-567 (Springer, Heidelberg, 1983).
11. Bretherton, F. P., Davis, R. E. & Fandry, C. B. *Deep Sea Res.* 23, 559-582 (1976).
12. Carter, E. F. & Robinson, A. R. *J. atmos. ocean. Technol.* (in the press).
13. Pedlosky, J. *Geophysical Fluid Dynamics* (Springer, Heidelberg, 1979).
14. Taupier-Letage, I. & Millot, C. *Oceanologica Acta* (in the press).
15. Robinson, A. R., Feliks, Y. & Pinardi, N. in *Proc. Lerici Workshop*, Lerici, Italy 1984 (ed. Charnok, H.) (in preparation).
16. Le Groupe Tourbillon *Deep Sea Res.* 30, 475-511 (1983).

Temporal changes in microseismicity and creep near Parkfield, California

C. M. Poley, A. G. Lindh, W. H. Bakun & S. S. Schulz

US Geological Survey, 345 Middlefield Road, MS 977, Menlo Park, California 94025, USA

The 25-km-long section of the San Andreas fault near Parkfield, California ruptured in similar magnitude-6 earthquakes in 1881, 1901, 1922, 1934 and 1966. On the basis of a number of lines of seismological evidence, a section of the San Andreas fault, now termed the Parkfield preparation zone, has been identified as the locus of the next Parkfield earthquake¹⁻³. Here we describe coincident changes in surface creep rates and deep seismicity near the Parkfield preparation zone following the 2 May 1983 Coalinga earthquake, and suggest that both respond to the same stimuli. These changes were concentrated near the point of initiation of the magnitude-6 characteristic Parkfield earthquakes, lending credence to the hypothesis that this section of the San Andreas fault is characterized by a unique set of physical properties⁴⁻⁶ which make it ideal for earthquake prediction studies and which may also be useful for identifying hypocentral regions in other areas.

Epicentres near the trace of the San Andreas fault (A-A' in Fig. 1) conform to a relatively simple pattern. Earthquakes between points A and B in Fig. 2a, located north-west of the Parkfield preparation zone (PPZ, Fig. 1), occur primarily between depths of 4 and 7 km, and correspond to the section of the San Andreas fault that is creeping at about 23 mm per yr⁷. Earthquakes between B' and A' in Figs 1 and 2a, south-east of the PPZ, tend to be deeper; their distribution resembles that of the aftershocks observed after the last characteristic Parkfield earthquake in 1966⁸. We have divided the area within the PPZ

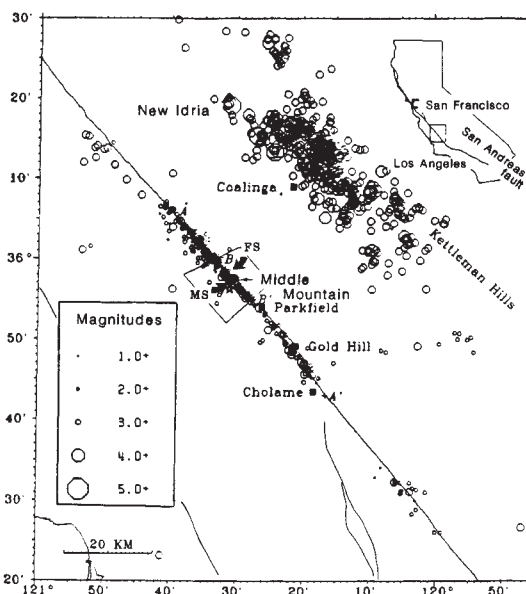


Fig. 1 Earthquake epicentres (1975-1985) relative to the trace of the San Andreas fault (long diagonal line) and the epicentres of the 1966 $M = 5.1$ foreshock and mainshock, shown as FS and MS respectively. Epicentre clusters east of the San Andreas fault are aftershocks of the 1975 Cantua Creek, 1976 Avenal, 1982 New Idria, 1983 Coalinga and 1985 Kettleman Hills earthquakes. Epicentres for all $M \geq 1.8$ earthquakes are shown, except the numerous $M \leq 3$ aftershocks of the 1983 Coalinga earthquake, which cover the Coalinga area when plotted. The Middle Mountain MM3 zone (quadrilateral) includes the preparation zone of the characteristic Parkfield earthquake.

Fig. 2 Cross-sections showing seismicity in the Parkfield area for 1975–1985. Symbol size is proportional to magnitude; the smallest symbol represents 1.8–2.29 in *a* and 1.5–2.29 in *b*. *a*, Seismicity of $M \geq 1.8$ along section A–A' (Fig. 1). Relative focal depths are generally accurate to less than 1 km; depths of the shallow shocks to the north-west of Middle Mountain have an uncertainty of ~ 2 km. 1966 aftershocks are shown as plus signs. Creepmeter locations are shown by four-letter names at the top. *b*, Seismicity of $M \geq 1.5$ along section B–B' (Fig. 1). Because hypocentre locations in *b* are based on a revised set of station travel-time corrections, hypocentres in *a* and *b* may differ for the same event. The hypocentres of the 1966 fore-shock and mainshock, and two 1975 $M = 4.4$ earthquakes are shown as stars labelled A, B, C and D respectively. A cluster of recurring $M = 3$ earthquakes is labelled E.

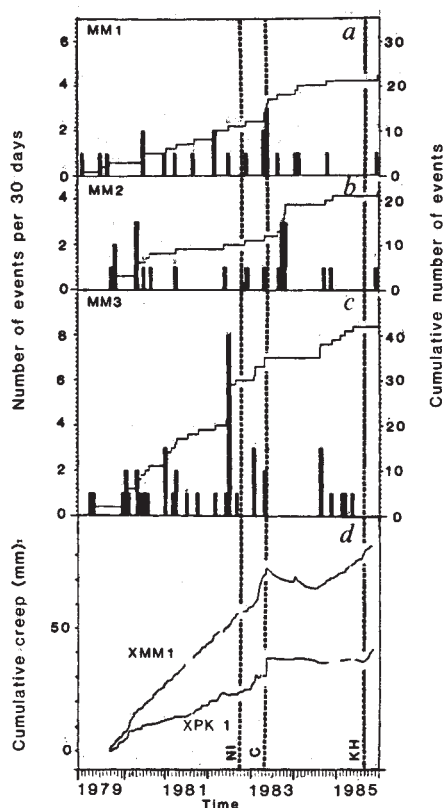
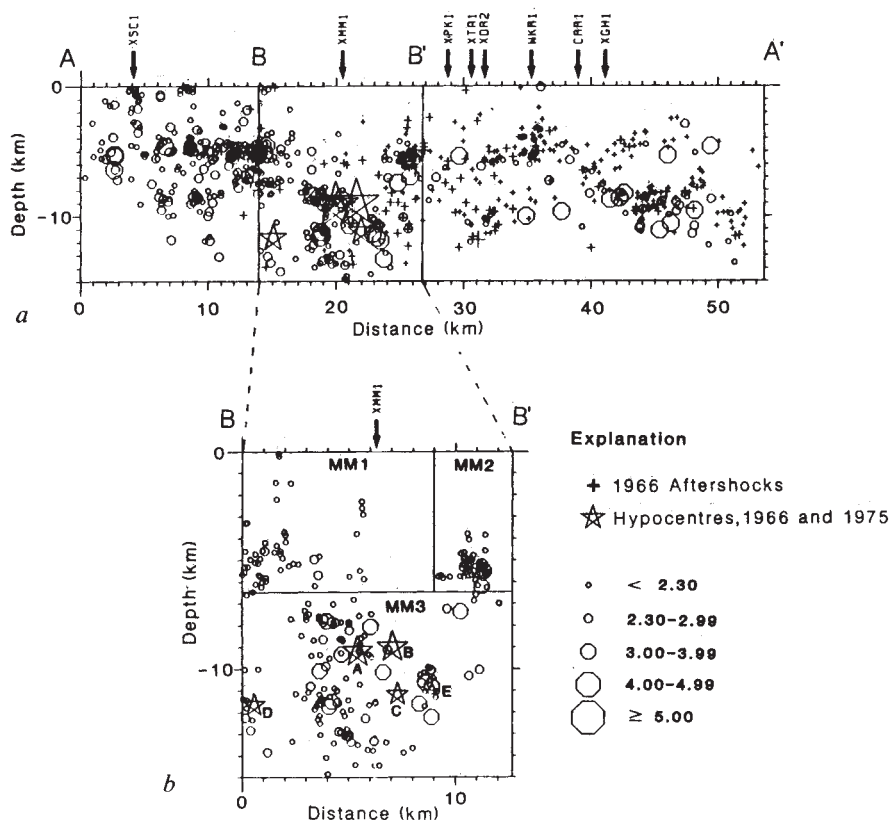


Fig. 3 Time histograms of seismicity $M \geq 1.6$ in: *a*, MM1; *b*, MM2; *c*, MM3 (Fig. 2*b*); and *d*, cumulative creep at creepmeters XMM1 and XPK1 (Fig. 2*a*). In the seismicity figures (*a*, *b* and *c*) the continuous solid line shows the cumulative number of events (label on right axis) and the bar graphs show the number of events per 30-day period (left axis). Times of the New Idria (NI), Coalinga (C) and Kettleman Hills (KH) main shocks are shown by vertical dashed lines.

(B to B' in Fig. 2*b*) into three parts on the basis of the distribution of seismic activity. (1) MM1 (in Fig. 2*b*) has the shallow seismicity in the north-west that is the continuation of the shallow activity farther to the north-west; no events of $M \geq 4$ have occurred in MM1 since 1969. (2) MM2 includes the cluster of shallow events within the inferred rupture zone⁶ of the Parkfield main shocks. (3) MM3, the deep section beneath Middle Mountain, includes the hypocentre of the 1966 earthquake⁹ and shows the most activity.

The seismicity within MM3 (Fig. 3) is remarkable in several respects. First, the deepest microearthquake locations along this section of the San Andreas occur within the lower half of MM3 (Fig. 2). Other clusters of deep seismicity have been associated with potential initiation points. Moths *et al.*¹⁰ identified one such cluster north of San Juan Bautista, 140 km to the north-west of Parkfield along the San Andreas. Olson¹¹ described a deep cluster near San Francisco, which may in fact coincide with the epicentre of the 1906 San Francisco earthquake¹². We do not have sufficient information, however, to conclude that this is a general feature of initiation points. Second, the *b*-slope is about 0.6, the lowest estimate for *b*-slope in central California¹³. Third, hypocentres of most of the $M \geq 3$ shocks in the Parkfield area since 1970 are within MM3. The two largest shocks in the Parkfield area since 1966 occurred in MM3 in 1975, $M = 4.4$ and $M = 4.8$ events at 11–12 km depth (smaller stars labelled C and D in Fig. 2*b*). Fourth, most of the other earthquakes within MM3 occur in spatial clusters, each of which has been the locus of repeated episodes of seismic activity since 1970. Cluster E (Fig. 2*b*) is characterized by a remarkable sequence of $M \geq 3$ events at regular 39- to 40-month intervals since 1967¹⁴. Finally, cluster E has also been identified as a source of shocks having larger stress drops than those on the adjacent parts of MM3¹⁵. These observations taken together strongly suggest that the hypocentral region of the characteristic Parkfield earthquake (the characteristic hypocentre) is seismologically unique along the San Andreas fault, so that it might be identifiable, even if we did not have detailed records from the last two major sequences to pinpoint it. We do not believe that the evidence is sufficient to conclude that this suite of features is a general feature of

initiation points, along the San Andreas fault or elsewhere; we do believe that similar characteristics might be worth looking for elsewhere.

In an effort to identify anomalous, and possibly precursory, seismicity within the preparation zone, we have examined the temporal patterns of the microseismicity near Parkfield. The seismicity rate for events larger than $M = 1.75$ has been relatively constant since 1970, except for a temporary increase in 1975 associated with two $M = 4.4+$ Middle Mountain events and the 15-month hiatus following the 2 May 1983, $M = 6.7$ Coalinga earthquake. The decrease in seismicity following the Coalinga main shock is particularly clear within MM3, where activity ceased entirely at the $M \geq 1.6$ level at the time of the Coalinga earthquake, remained quiet for 15 months, and resumed in August 1984 (Fig. 3c).

Ten creepmeters in the Parkfield area are used to monitor surface slip across the fault (Fig. 2a). Measured creep rates decrease from north to south, from 23 mm per yr at Slack Canyon (XSC1) to 12 mm per yr at Parkfield (XPK1) to 3.3 mm per yr at Gold Hill (XGH1)⁷. Beginning in January 1983, small surges in creep rate were observed at XMM1 and XPK1. Following the co-seismic slip at the time of the Coalinga shock, creep rates along the Parkfield section were lower than pre-1983 rates⁴ (Fig. 3d) and left-lateral displacement was recorded at XMM1 and XPK1, although it was less pronounced at XPK1. Right-lateral slip resumed at XMM1 in August 1984, coincident with the resumption of seismicity in MM3. Right-lateral motion resumed at XPK1 a year later, shortly after the 4 August 1985, $M = 5.7$ earthquake beneath the north end of the Kettleman Hills anticline¹⁶ (Figs 1 and 3). Thus, fault creep near the PPZ appears to be associated not only with seismicity along the San Andreas fault, but also with the small stress perturbations caused by moderate to large off-fault earthquakes.

The hypothesis of a time-stationary PPZ is based primarily on the facts that historic Parkfield magnitude-6 shocks have tended to initiate at the same point, and had energetic foreshock sequences there in 1934 and 1966 as well. Taken together with the additional characteristics of the seismicity and creep of the PPZ that were described above, we believe these data are most easily explained by postulating that the PPZ differs in some physical or mechanical property from the surrounding fault.

This same PPZ section of the fault was the most strongly affected by the Coalinga earthquake, even though given the distance to the Coalinga source region one would have expected the stress perturbation to be rather more evenly distributed along the fault¹⁷. Using the dislocation program of Mavko *et al.*¹⁷ to model the Coalinga earthquake, R. W. Simpson (personal communication) has calculated the static horizontal shear stress changes across the San Andreas fault. The calculated near-surface changes are negative (that is, they retard right-lateral slip) through the Parkfield region, which agrees with the long-term retardations observed at most of the creepmeters in the area⁴. The calculated stress changes are broadly peaked, with the centre of the peak within the PPZ, which agrees with the large changes seen at XMM1 and XPK1 (Fig. 3). The predicted stress changes are more smoothly distributed along the fault however, than the observed creep changes which peaked very sharply in the PPZ. From this and from the coincident shut-off of seismicity in the MM3 area directly below XMM1 (Fig. 4), we infer that the PPZ is much more sensitive to small stress perturbations than are the adjacent sections of the fault.

If we carry this assumption one step further, we can speculate that the enhanced stress sensitivity on and beneath Middle Mountain provides a simple physical mechanism to explain some of the other phenomena that characterize the PPZ. Theoretical and laboratory analyses suggest that the occurrence of a large-scale mechanical instability on a fault within the Earth (an earthquake) requires some form of strain (or strain-rate) softening; that is, as a fault approaches failure, it begins to weaken as the amount (or rate) of deformation increases^{18,19}.

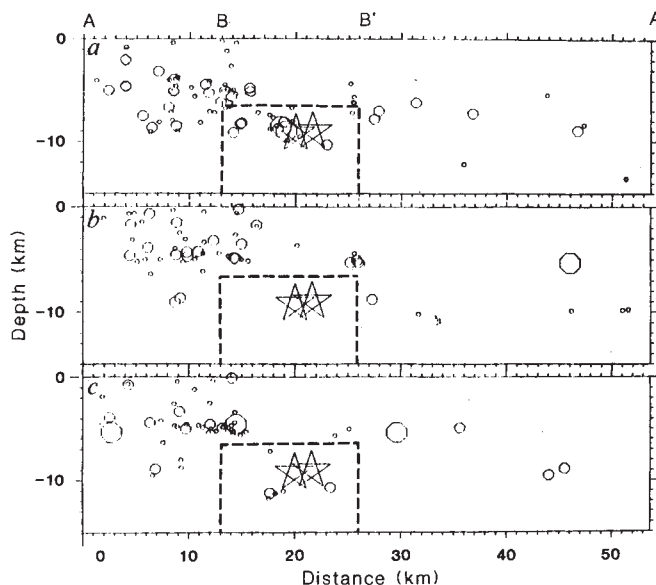


Fig. 4 Cross-sections showing short-term seismicity $M \geq 1.6$ in the Parkfield area. Symbols are the same as in Fig. 2. Dashed boxes show area of MM3. a, Seismicity during the 14 months before the 2 May 1983 Coalinga earthquake (1 March 1982–1 May 1983). b, Seismicity during the 15-month quiet period following the Coalinga earthquake (2 May 1983–31 July 1984). c, Seismicity from 1 August 1984–31 December 1985.

It seems reasonable to suppose that a repeated hypocentre marks a point on a fault at which the stress-strain relation takes some extreme value. We speculate that a region such as the PPZ, where significant shocks start and which is characterized by anomalous seismicity and creep, might show more extreme strain (or strain-rate) softening behaviour than would the adjacent sections of the fault. We suggest that the repeated occurrence of Parkfield foreshocks and mainshock hypocentres within the PPZ is due to the same mechanism that caused the changes following the Coalinga earthquake to be larger within the PPZ than along other parts of the Parkfield section of the San Andreas fault.

We further suggest that any deformation premonitory to the next Parkfield shock might be concentrated beneath Middle Mountain in the identified region of enhanced stress sensitivity within the PPZ, and efforts to detect such deformation should be concentrated there. We should add however, that we do not have any basis on which to assume that the PPZ is 'stronger' or 'weaker' than the surrounding fault zone or whether it should be considered a 'barrier' or an 'asperity'. The high-stress drops from cluster E suggest that at least that point might be stronger; but attempts to model the Parkfield region suggest that the highest yield strengths are located 10 km to the south-east^{18,19}. Thus we only wish to note the number of remarkable characteristics clustered around this small fault patch. The term 'enhanced stress sensitivity' to describe the PPZ seems to us adequately descriptive and appropriately vague. We do believe that the evidence for enhanced stress sensitivity beneath Middle Mountain provides a basis on which to begin to seek a physically plausible unifying concept for earthquake prediction.

Thorough and critical reviews from Fred Klein and Paul Reasenberg, and enlightening discussions with Bob Simpson and Jim Dieterich, greatly enhanced and substantially improved the manuscript.

Received 28 October 1986; accepted 5 March 1987.

1. Bakun, W. H. & McEvilly, T. V. *J. geophys. Res.* **89**, 3051–3058 (1984).
2. Bakun, W. H. & Lindh, A. G. *Earthquake Pred. Res.* **3**, 285–304 (1985).
3. Shearer, C. F. *U.S. geol. Surv. Open-File Report* 85–201 (1985).
4. Schulz, S. S., Mavko, G. M. & Brown, B. D. *Prof. Pap. U.S. geol. Surv.* (in the press).
5. King, G. *Pure appl. Geophys.* **121**, 761–815 (1983).
6. Lindh, A. G. & Boore, D. M. *Bull. seism. Soc. Am.* **71**, 95–116 (1981).

7. Schulz, S. S., Mavko, G. M., Burford, R. O. & Stuart, W. D. *J. geophys. Res.* **87**, B6977-B6982 (1982).
8. Eaton, J. P., O'Neill, M. E. & Murdock, J. N. *Bull. seism. Soc. Am.* **60**, 1151-1197 (1970).
9. Bakun, W. H. & Lindh, A. G. *Science* **229**, 619-624 (1985).
10. Moths, B. L., Lindh, A. G., Ellsworth, W. L. & Fluty, L. *Eos* **62**, 958 (1981).
11. Olson, J. A. *Bull. R. Soc. N.Z.* **24**, 87-97 (1986).
12. Bolt, B. A. *Bull. seism. Soc. Am.* **50**, 457-471 (1968).
13. Lindh, A. G. & Jones, L. *Eos* **66**, 982 (1985).
14. Poley, C. M., Lindh, A. G. & Bakun, W. H. *Earthquake Not.* **57**, 21 (1986).
15. O'Neill, M. E. *Bull. seism. Soc. Am.* **74**, 27-40 (1984).
16. Schulz, S. S. & Burford, R. O. *Eos* **66**, 985 (1985).
17. Mavko, G., Schulz, S. S. & Brown, B. D. *Bull. seism. Soc. Am.* **75**, 475-489 (1985).
18. Stuart, W. D., Archuleta, R. J. & Lindh, A. G. *J. geophys. Res.* **90**, 595-604 (1985).
19. Tse, S. T., Dmowska, R. & Rice, J. R. *Bull. seism. Soc. Am.* **75**, 709-736 (1985).

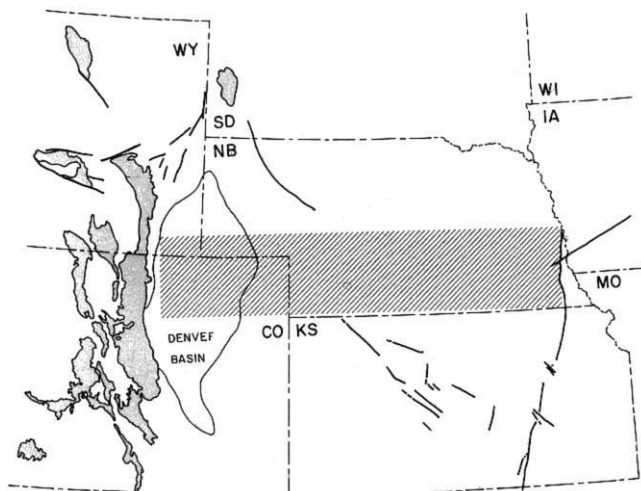


Fig. 1 Simplified tectonic map of the central United States³², including the Great Plains, Denver Basin and Precambrian uplifts of the Rocky Mountains. The study area is marked by the shaded rectangle. Bouguer gravity, topography, and basement data along four evenly spaced profiles are compiled in Fig. 2.

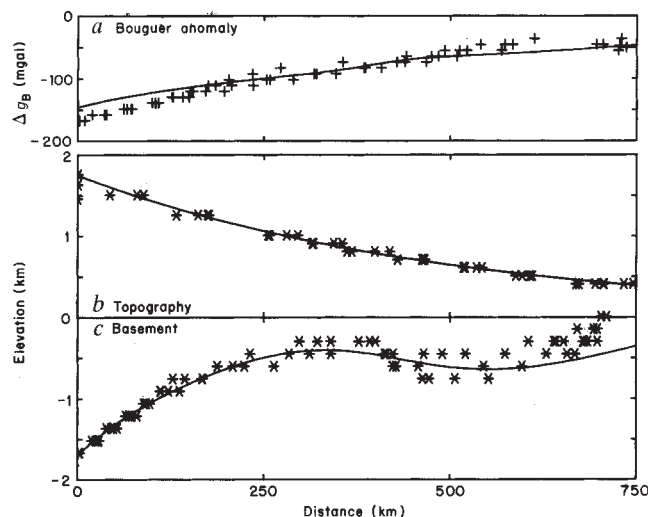


Fig. 2 *a*, Bouguer gravity anomaly³³ Δg_B over the study area (wavelengths of less than 250 km have been removed). Anomaly predicted by flexural and gravity models is given by the solid curve. *b*, Surface elevations from a 5- and 5-min averaged topography map³⁴. Solid curve shows that elevations decrease exponentially to the east with a length scale of 500 km. *c*, Depth to Precambrian basement³². Basement depths predicted by flexural modelling are given by the solid curve.

Buoyant sub-surface loading of the lithosphere in the Great Plains foreland basin

Charles L. Angevine & Kathryn M. Flanagan

Department of Geology and Geophysics, University of Wyoming, Laramie, Wyoming 82071, USA

A recent development¹ in the study of foreland sedimentary basins is the recognition of the important role played by negatively buoyant sub-surface loads in the evolution of some of these basins. Models² for the subsidence of the Apennine and Carpathian foreland basins, in particular, show that sub-surface loads can equal or exceed the topographic load of the neighbouring mountain belt. Here we show that part of the Rocky Mountain foreland is underlain by a positively buoyant sub-surface load which counterbalances almost half of the total topographic load on the western Great Plains. The presence of this load explains the shallowness of the Denver Basin and the large Bouguer gravity anomalies over the Great Plains. We suggest that this load was emplaced during the Miocene when, according to geological evidence³⁻⁵, the region underwent epeirogenic uplift. The sub-surface load may result from lithospheric thinning, which often accompanies or precedes crustal extension^{6,7}.

The study area extends from the axis of the Denver Basin in the west to the Missouri River in the east, a distance of almost 750 km (Fig. 1). Over this distance, the Great Plains rise smoothly from an elevation of 400 m in the east to almost 2,000 m along the Front Range of the Rocky Mountains (Fig. 2*b*). Basement profiles are considerably more complicated (Fig. 2*c*). The Denver Basin, immediately adjacent to the Rockies, contains up to 4 km of Phanerozoic sediments, and is bordered by the Chadron-Cambridge arch to the north-east. Although the asymmetric shape of the Denver Basin is consistent with flexure of the lithosphere⁸⁻¹¹ under the load of the Rockies and topography of the Great Plains, the basin is anomalously shallow; simple isostatic considerations suggest that it should contain more than 6 km of sediment. Thus, the Denver Basin cannot be explained solely by flexure of the lithosphere under the load of the present topography.

Sub-surface mass anomalies, which may also be involved in the deflection of the lithosphere, can be detected using gravity anomalies¹. The Bouguer anomaly, which becomes increasingly negative towards the Rockies (Fig. 2*a*), is largely associated with the deflection of basement and the Moho¹². If we assume that the crust is of uniform thickness and that its surface lay at sea-level before loading, this contribution can be removed from the Bouguer anomaly to obtain the residual gravity anomaly shown in Fig. 3*a*. In making this reduction we assumed a density contrast of 800 kg m^{-3} between mantle and sediments. With a maximum amplitude of -110 mgal the residual anomaly is quite large and suggests that a mass deficiency lies beneath the Great Plains. Even if the basement lay 300 m below sea-level before

loading our reduction technique would introduce a false anomaly of only -10 mgal . Therefore, we are confident that the residual anomaly is real and that a buoyant root exists. The long wavelength of the gravity anomaly ensures that the magnitude of the corresponding mass anomaly will not be a strong function of depth. One possible model (Fig. 3*b*) has a magnitude of $2.2 \times 10^6 \text{ kg m}^{-2}$ under the western Plains, but decreases to $7.2 \times 10^5 \text{ kg m}^{-2}$ near the Chadron-Cambridge Arch. If located at a depth of 80 km this mass anomaly would generate the gravity anomaly shown by the solid curve in Fig. 3*a*. Agreement between model and residual anomalies is quite good except near the Rockies, where the misfit may be associated with an additional mass deficiency beneath the Front Range. Because the magnitude of the mass anomaly is essentially independent of the depth of the anomaly, we could place the mass anomaly at any reason-

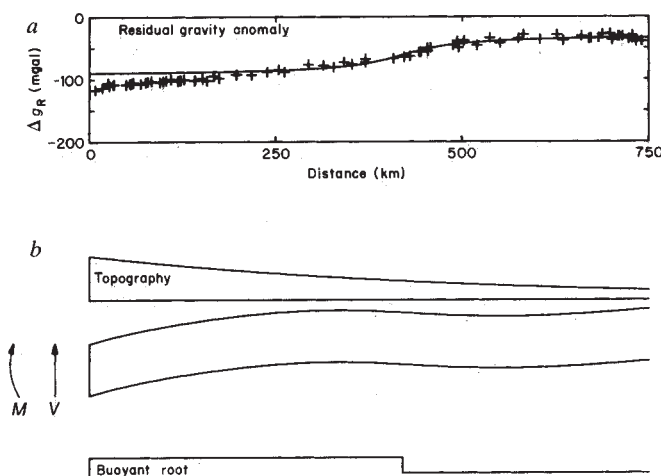


Fig. 3 *a*, Residual gravity anomaly, Δg_R , calculated by subtracting gravity anomalies caused by the deflection of basement and Moho from the Bouguer anomaly. This residual anomaly shows that the central Great Plains are underlain by a mass deficiency. A solid curve shows the gravity for a model anomaly, given in *b*, located at a depth of 80 km. The long wavelength of the mass anomaly ensures that a good fit to the gravity anomaly will be obtained, no matter how deep (or shallow) the anomaly is placed. *b*, A schematic view of the loads acting upon the elastic plate. The load of the topography is partially compensated by the upward push of the buoyant mass anomaly. A bending moment, M , and shear force, V , at the end of the plate represent the load imparted by the Rocky Mountains. The sense of the bending moment suggests that the plate continues under the Rockies, whereas the shear force indicates that the Rockies are largely supported from below by a buoyant mass anomaly.

able depth and obtain an equally good fit to the residual gravity anomaly.

We model subsidence of the lithosphere as the two-dimensional flexure of a semi-infinite elastic plate¹⁰ under distributed and end loads (Fig. 3*b*). The distributed loads are due to topography and the low-density root, whereas the bending moment and shear force applied at the end of the plate represent the load imparted by the Rocky Mountains. Both end loads and the flexural rigidity of the lithosphere are unknowns which must be determined. In practice, the flexural profile is constrained to match basement depth and slope in the Denver Basin and the flexural rigidity is adjusted to obtain the best fit over the rest of the profile. This procedure is possible only because the magnitude of the sub-surface load is well-determined from the gravity analysis. An excellent fit (Fig. 2*c*) to the basement deflection is obtained using a flexural rigidity of 10^{24} Nm, a bending moment of 6.8×10^{16} Nm, and a shear force of 7.8×10^{11} N. The quality of the fit deteriorates markedly if the flexural rigidity is increased or decreased by an order of magnitude. Although the sense of the bending moment is consistent with the extension of the elastic lithosphere under the Rockies, the direction of the shear force suggests that a large fraction of topography in the Rocky Mountain province is also supported by a low-density root located beneath, and pushing upward on, the lithosphere. If the Rockies were simply a surficial load pushing downward on the lithosphere the shear force would act in the opposite direction. The Bouguer anomaly, calculated from the flexural profile, also agrees well with the observed anomaly (Fig. 2*a*).

Emplacement of the sub-surface load and the resultant epeirogenic uplift of the Rocky Mountains and Great Plains may have occurred late in Cenozoic history. The Great Plains were covered by an inland sea during the late Cretaceous¹³ and well into the early Tertiary^{14,15}. Uplift of the Front Range in latest Cretaceous through early Eocene time was marked by deposition of the Laramie Formation and Dawson Arkose in

the Denver Basin⁵. However, the scarcity of syntectonic sediments suggests that the mountains did not possess enough relief to generate a foredeep basin of significant depth during this time¹⁶. The mid to late Eocene history on the Plains is one of erosion or nondeposition, with palaeosol formation on Dawson Arkose and late Cretaceous Pierre Shale^{5,17}. Oligocene and early Miocene sediments^{4,18} in eastern Colorado and Nebraska record renewed uplift in the Rockies and concurrent volcanism. A change in sedimentation begins sometime during the mid-Miocene, marked by extensive downcutting in Nebraska⁴. This pattern continues into the later Miocene and Pliocene, stripping Neogene sediments from north-central Colorado. This major change in depositional regime occurs on a regional scale suggesting that emplacement of the sub-surface load and the resulting regional uplift of the Rockies and Great Plains began during the Miocene.

More than ten years ago, Suppe *et al.*⁶ suggested that the late Cenozoic uplift of the western United States is associated with the erosion of the underlying lithosphere by mantle plumes and that the uplift is a precursor to Basin and Range style crustal extension. De-lamination^{19,20} and stoping⁷ may be alternative ways to thin the mantle lithosphere. If the density contrast between the mantle lithosphere and asthenosphere is 1.5% then the sub-surface load can be generated by thinning the lithosphere by 15 km in the east and 45 km in the west. Delay times for P-wave arrivals^{21,22} on the Great Plains are consistent with a decrease in lithospheric thickness to the west. The lack of a significant heat flow anomaly²³ on the western Plains is also consistent with a thinning event in the Miocene, because the thermal perturbation would not have had sufficient time to reach the surface²⁴. Lithospheric thinning should generate large extensional stresses^{7,25} and could lead to normal faulting of the crust above the thinned region. Normal faults of Miocene age are observed in both Colorado and Wyoming²⁶. Thus, there is ample reason to associate the sub-surface load with lithospheric thinning under the Rocky Mountains and Great Plains. We note, however, that the sub-surface load is responsible for <900 m of total elevation on the western Plains. The balance of the elevation results from the fall of sea level since the Cretaceous and, perhaps, the removal of a sub-surface load^{27,28} which caused anomalous subsidence in the region during the latest Cretaceous and Palaeocene.

The prominence of the Chadron-Cambridge Arch is due largely to the geometry of the sub-surface mass anomaly (Fig. 3*b*). However, the patterns of erosion and sediment dispersal in the area¹⁸ suggest that the arch formed before the Oligocene. We believe that a portion of the sub-surface mass anomaly, near the Chadron-Cambridge Arch, must be associated with crustal thickening resulting from horizontal compression during the latest Cretaceous through to the Eocene. Thus the sub-crustal anomaly tapers to the east more smoothly than indicated.

The presence of a buoyant mass anomaly beneath the Rocky Mountains and Great Plains appears to be related to the active extension of crust in the western United States. As such, the results of this study may be applicable in other continental back-arc regions undergoing extension. A recent study²⁹ of the Bolivian Andes suggests that the topographic load of the sub-Andes and Cordillera Oriental is too large to explain the deflection of the Brazilian Shield as determined from gravity anomalies. There is active extension in the high Andes at present³⁰ and some evidence³¹ for a thin lithosphere in the region. Perhaps some portion of the topography in this area is supported by sub-surface loads. This possibility could be evaluated through combined flexural and gravity modelling if there was better control on the depth to basement under the Chaco Plain and sub-Andes.

We thank D. Sahagian for providing a preprint of his paper on epeirogeny in the western United States and R. W. Simpson for supplying the topographic data used here. We have benefitted from discussions with S. J. Babits, J. A. Lillegraven and J. D.

Love. This research was supported in part by the Donors to the Geophysics Development Fund and a University of Wyoming A & S Basic Research Grant to C.L.A.

Received 15 September 1986; accepted 11 February 1987.

1. Karner, G. D. & Watts, A. B. *J. geophys. Res.* **88**, 10449-10477 (1983).
2. Royden, L. & Karner, G. D. *Nature* **309**, 142-144 (1984).
3. Stanley, K. O. *Geol. Soc. Am. Bull.* **87**, 297-309 (1976).
4. Trimble, D. E. *Mount. Geologist* **17**, 59-69 (1980).
5. Swinehart, J. B., Souders, V. L., DeGraw, H. M. & Diffendal, R. F. in *Cenozoic Paleogeography of the West-Central United States* (eds Flores, R. M. & Kaplan, S. S.) 209-229 (SEPM, Denver, 1985).
6. Suppe, J., Powell, C. & Berry, R. *Am. J. Sci.* **275-A**, 397-436 (1975).
7. Turcotte, D. L. *J. geol. Soc. Lond.* **140**, 701-724 (1983).
8. Beaumont, C. *Geophys. J. R. astr. Soc.* **65**, 291-329 (1981).
9. Jordan, T. E. *Bull. Am. Ass. Petrol. Geol.* **65**, 2506-2520 (1981).
10. Turcotte, D. L. & Schubert, G. *Geodynamics* (Wiley, New York, 1982).
11. Lyon-Caen, H. & Molnar, P. *J. geophys. Res.* **88**, 8171-8191 (1983).
12. Vening Meinesz, F. A. *Proc. ned Akad. Wet.* **44**, 1-12 (1941).
13. King, P. B. *The Evolution of North America* (Princeton University Press, 1977).
14. Cherven, V. B. & Jacob, A. F. in *Cenozoic Paleogeography of West-Central United States* (eds Flores, R. M. & Kaplan, S. S.) 127-170 (SEPM, Denver, 1985).
15. Marinovich, L., Brouwers, E. M. & Carter, L. D. *Geology* **13**, 770-773 (1985).
16. Morse, D. G. in *Cenozoic Paleogeography of the West-Central United States* (eds Flores, R. M. & Kaplan, S. S.) 277-292 (SEPM, Denver, 1985).
17. Soister, P. E. & Tschudy, R. H. in *Energy Resources of the Denver Basin* (eds Pruitt, J. & Coffin, P. E.) 231-235 (RMAG, Denver, 1985).
18. Stanley, K. O. & Wayne, W. J. *Geol. Soc. Am. Bull.* **83**, 3675-3690 (1972).
19. Bird, P. *J. geophys. Res.* **84**, 7561-7571 (1979).
20. Bird, P. *Tectonics* **3**, 741-758 (1984).
21. Cleary, J. & Hales, A. L. *Bull. seism. Soc. Am.* **56**, 467-489 (1966).
22. Herrin, E. & Taggart, J. *Bull. seism. Soc. Am.* **58**, 1325-1337 (1968).
23. Lachenbruch, A. H. & Sass, J. H. in *The Earth's Crust* (ed. Heacock, J. G.) 626-675 (Geophys. Monogr. Ser. 20, Am. Geophys. Union, Washington, DC, 1977).
24. Sahagian, D. *J. geophys. Res.* (in the press).
25. Morgan, P. & Swanberg, C. A. *J. Geodynamics* **3**, 39-63 (1985).
26. Izett, G. A. in *Cenozoic History of the Southern Rocky Mountains* (ed. Curtis, B. F.) 179-210 (Memoir 144, Geol. Soc. Am., 1973).
27. Cross, T. A. & Pilger, R. H. *Nature* **274**, 653-657 (1978).
28. Beaumont, C. *Eleventh International Congress on Sedimentology* (eds Nriagu, J. O. & Troost, R.) (McMaster University, Hamilton, 1982).
29. Lyon-Caen, H., Molnar, P. & Suarez, G. *Earth planet. Sci. Lett.* **75**, 81-92 (1985).
30. Mercier, J. L. *Phil. Trans. R. Soc. A300*, 337-355 (1981).
31. Froidevaux, C. & Isacks, B. L. *Earth planet. Sci. Lett.* **71**, 305-314 (1985).
32. King, P. B. *Tectonic Map of North America* (US Geol. Surv., Reston, 1976).
33. Hildebrand, T. G., Simpson, R. W., Godson, R. H. & Kane, M. F. *Digital Colored Residual and Regional Bouguer Gravity Maps of the Conterminous United States with Cut-Off Wavelengths of 250 km and 1000 km* (US Geol. Surv., Reston, 1982).
34. Simpson, R. W., Jachens, R. C., Blakely, R. J. & Saltus, R. W. *J. geophys. Res.* **91**, 8348-8372 (1986).

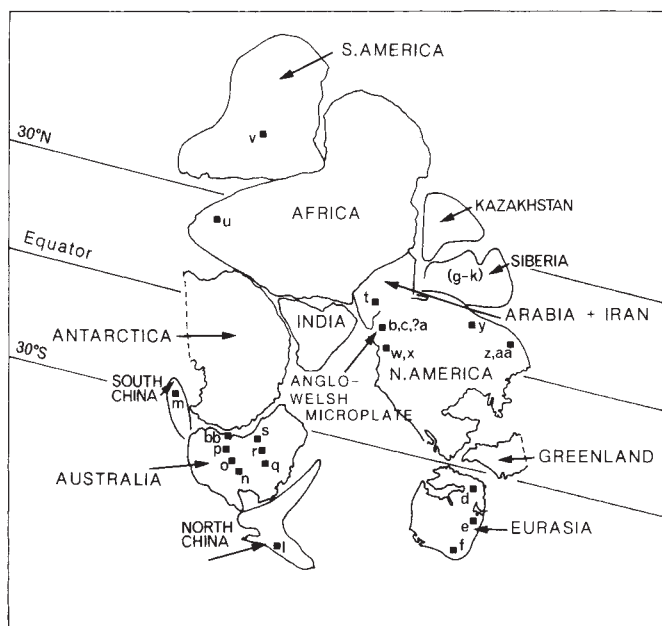


Fig. 1 Sketch map (approximately cylindrical equidistant projection) of the late Precambrian super-continent of Piper^{3,4}. Key to Ediacaran fossil localities^{1,11-15}: a, south-east Newfoundland; b, Carmarthen, Wales; c, Charnwood, England; d, Lake Torneträsk, Sweden; e, White Sea, N USSR; f, Black Sea, south-west Ukraine; g, W Urals; h, Yenisey River, USSR; i, Anabar and Olenek, N Siberia; j, River Maya, E Siberia; k, Lake Baikal, S Siberia; l, Longshan, NE China; m, Yangtze Gorge; n, SW Georgina Basin, Australia; o, Deep Well, SE of Alice Springs; p, Ediacara, S Australia; q, Mt Skinner, Australia; r, Laura Creek, SW of Alice Springs; s, Punkerri Hills, S. Australia; t, central Iran; u, Namibia, SW Africa; v, Mato Grosso, SW Brazil; w, Durham, N Carolina; x, Stanley County, N Carolina; y, Rocky Mountains, British Columbia; z, Mackenzie Mountains, NW Canada; aa, Wernecke Mountains, Yukon; bb, Adelaide geosyncline, S Australia. Placing of some localities approximate, particularly where crowded.

The fit of the continents in the late Precambrian

Stephen K. Donovan

Department of Geology, University of the West Indies, Mona, Kingston 7, Jamaica

The relative positions of the continents during the late Precambrian (~600 Myr BP) are still largely a matter of speculation¹. Three reconstructions²⁻⁵, based principally on palaeomagnetic and tectonic data, suggest that the continents were joined to form one supercontinent. This supercontinent may have been approximately equatorial², subpolar to subpolar^{3,4}, or polar to equatorial⁵ in trend. Here I test these reconstructions by plotting the distribution of the late Precambrian Ediacaran fauna¹ and show that the equatorial and subpolar to subpolar models are the most acceptable as templates for metazoan palaeobiogeography.

Reconstruction of the relative positions of the ancient continents relies on a number of criteria, particularly the linear magnetic anomalies of the sea floor, palaeomagnetism of continental rocks, relationships of facies, structures and mineralization episodes between continents, and analysis of the palaeobiogeographic distribution of the ancient organisms. The oldest sea floor is Jurassic in age, so reconstruction of pre-Jurassic plate orientations relies heavily upon criteria other than linear magnetic anomalies. Phanerozoic palaeocontinental world maps have now been constructed on the basis of

palaeontological and palaeomagnetic evidence^{6,7}, the most convincing restorations being produced where the two approaches have been combined^{8,9}. However, in contrast to the relatively common palaeocontinental restorations of the Phanerozoic, only three reconstructions have been published for the late Precambrian (Ediacaran)²⁻⁵. These models group the global landmass together as one supercontinent and should be considered as a basis for further testing rather than final solutions¹. This paper is not a discussion of the validity of the palaeomagnetic and tectonic data used in each reconstruction but rather an attempt to apply the same geological evidence to all three models to assess their validity. I have attempted to test all three restorations on the basis of available palaeobiogeographic data, following recent publication of important reviews of the entire late Precambrian (Ediacaran) metazoan fauna^{1,10}.

Late Precambrian metazoan faunas, dominated by cnidarians, are now known from over 25 localities worldwide^{1,10-15}. At present the available stratigraphic information does not permit close correlation between localities, although they are of at least approximately the same age, based on the occurrence of particular genera and even species at a number of outcrops (Table 1). For example, the medusoid ('jellyfish') *Cyclomedusa davidi* Sprigg has been reported from Ediacara in South Australia, Leicestershire in the United Kingdom, the Wernecke Mountains in the Yukon and the Rocky Mountains in British Columbia (but see comments by Weiguo¹⁶). Some localities have yielded only one or a few taxa, whereas in both the White Sea and Black Sea areas of the USSR perhaps 60 to 80 species are present¹. Further rich localities occur in Namibia and at Ediacara. Any statistical analysis of the similarity between different localities must wait until more taxa and faunas are adequately described

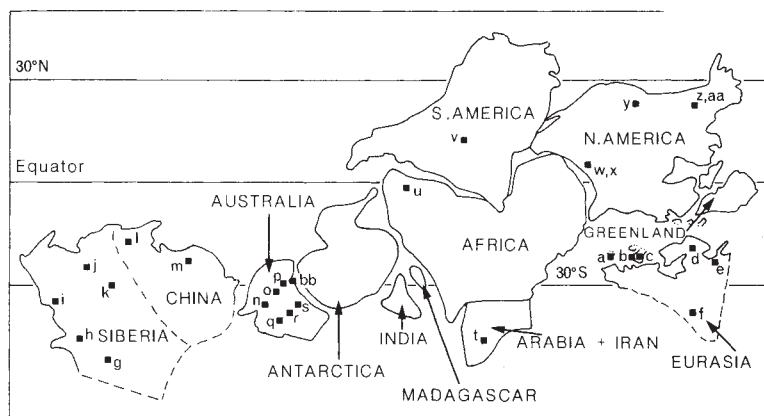


Fig. 2 Sketch map (approximately cylindrical equidistant projection) of the late Precambrian super-continent of Morel and Irving², showing distribution of fossiliferous localities. Locality information as for Fig. 1.

(for example, the rich fauna from Newfoundland has only been poorly interpreted^{17,18}). However, a speculative analysis has been attempted on the basis of available plate reconstructions.

The distribution of the 28 Ediacaran fossil localities on the three base maps (Figs 1–3) differs considerably. It is a general rule that shelf faunas at similar latitudes may be the same or different due to dispersal factors, whereas the shelf faunas of different latitudes, but similar longitudes, are different, reflecting continental separation and climatic factors¹⁹. This is particularly applicable to sessile organisms which are reliant upon ocean currents to distribute their larvae. Free-floating organisms such as 'jellyfish', common both today and in the Ediacaran, show a much broader distribution even at the specific level. For example, *Aurelia aurita* (Linnaeus), the common jellyfish, is considered to have a worldwide distribution²⁰. Recent reports have noted it from the Mediterranean Sea, Atlantic Ocean, English Channel, North Sea and Baltic Sea²¹, in Chesapeake Bay (F. Jeal, personal communication) and off Japan (S. M. Head, personal communication).

Piper's restoration of the late Precambrian super-continent has undergone considerable modification since it was first pro-

posed²². The distribution of the Ediacara fauna in the latest reconstruction^{3,4} shows some apparently anomalous distributions (Fig. 1, Table 1). *Tribrachidium* and the annelids *Dickinsonia* and *Spriggina* are found on opposite sides of the U-shaped southern sea (localities e, p). Arthropods are similarly distributed, as well as being found on the Anglo-Welsh microplate^{4,23} (localities c, e, p). It is these organisms, particularly the arthropods, that we might expect to show close provincial relationships^{7,24}. *Cyclomedusa* is widely distributed both within and outside the U-shaped sea region (localities a–c, f, j, k, p, u, y, aa). Other medusoids have a less cosmopolitan and more disparate distribution, for example, *Tirassiana* is found in the Black and White Sea regions, Siberia and Namibia (localities e, f, g, u). Colonial cnidaria also have a patchy distribution.

Morel and Irving's reconstruction² (Fig. 2; Table 1) does not clearly indicate the positions of Eurasia, Siberia and China. I have placed Eurasia in close association with North America but have left the Chinese and Siberian Shields adjacent to Australia, as originally figured. This gives an essentially 'two-centre' distribution to the rarer, benthic, non-cnidarian elements of the fauna, although both centres (Australia and Eurasia) have

Table 1 Distribution of Ediacaran fossils that occur at multiple localities

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	aa	bb
Medusoids																												
<i>Cyclomedusa</i>	+	+	+	–	–	+	–	–	–	+	+	–	–	–	–	+	–	–	–	–	+	–	–	–	+	–	+	–
<i>Irridinitus</i>	–	–	–	–	+	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	–	–	–
<i>Protodipleurosoma</i>	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–	+	–	–	–	–	–	–	–	–	–	+	–	–
<i>Kullingia</i>	–	–	cf	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	–
<i>Medusinites</i>	–	+	–	–	–	+	–	–	–	–	–	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–	+	–
<i>Pseudorhizostomites</i>	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	–
<i>Tirassiana</i>	–	–	–	–	+	+	+	–	–	–	–	–	–	–	–	–	–	–	–	–	+	–	–	–	–	–	–	–
Colonial Cnidaria																												
<i>Pteridinium</i>	–	–	–	–	+	–	–	–	–	+	–	–	–	–	–	+	–	–	–	–	+	–	–	–	+	–	–	+
<i>Inkrylovia</i>	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	+	–
<i>Charnia</i>	+	–	+	–	+	–	–	–	–	–	–	–	+	–	–	–	–	–	–	+	–	–	–	–	–	–	–	–
<i>Charniodiscus</i>	+	–	+	–	+	–	–	–	–	–	–	–	–	–	+	+	–	–	–	–	–	–	–	–	–	–	+	–
<i>Glaessnerina</i>	–	–	–	–	+	–	–	–	+	–	–	+	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	–
<i>Arumberia</i>	–	–	–	–	–	–	+	–	–	–	–	cf	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Annelids																												
<i>Cloudina</i>	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	+	–	–	–	–	–
<i>Dickinsonia</i>	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	–
<i>Spriggina</i>	–	–	–	–	cf	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Arthropods																												
<i>Vendomiids</i>	–	–	+	–	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
<i>Parvancorina</i>	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Phylum uncertain																												
<i>Tribrachidium</i>	–	?	–	–	+	–	–	–	–	–	–	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	–

Information from various sources^{1,10–15,23}. +, present; –, absent; ?, questionable occurrence; cf, a genus similar to, but not absolutely identical with, this taxon. The faunas of localities e and f (see key in caption to Fig. 1) are very similar¹. But both are still under investigation and complete faunal lists are not available at present. A number of taxa shown as absent in column f are therefore probably present.

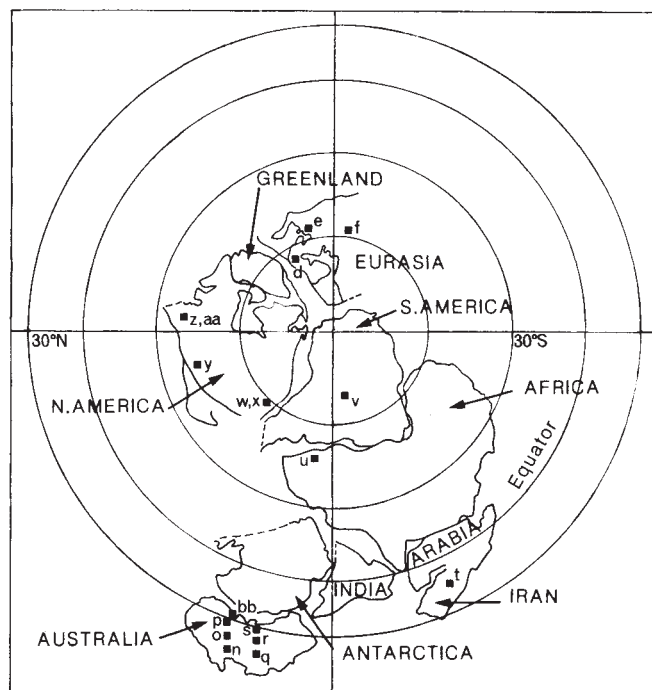


Fig. 3 Sketch map (south polar Lambert equal-area projection) of the late Precambrian supercontinent of Bond *et al.*⁵, showing distribution of fossiliferous localities. Locality information as for Fig. 1.

similar latitudes (compare with Piper's model). *Cyclomedusa* is limited to the Southern Hemisphere, apart from one North American occurrence. Other medusoids and colonial cnidaria have a similar distribution, many genera being known from the Southern Hemisphere and North America. *Charnia* is notable in only being found in the Southern Hemisphere (localities a, c, e, m, t).

The model of Bond *et al.*⁵ is incomplete, with no indication of the positions of the Siberian and Chinese Shields. *Tribrachidium*, *Dickinsonia*, *Spriggina* and the arthropods occur at two widely separated localities, about 30°N and 60°S. This is improbable.

A fourth restoration²⁵ only considers the relationship of western North America to China and Australia. No latitudinal data is provided and the faunas of these areas are not particularly complementary.

I conclude that the models of Piper^{3,4} and Morel and Irving² give the most probable palaeobiogeographic distributions of the Ediacara fauna. The rarer elements of the benthos, *Tribrachidium*, the arthropods and the annelids, are found in two main areas, Australia and the White Sea/Black Sea region. These lie at about 30°S in both reconstructions but are separated by either broad, enclosed sea (Fig. 1) or a long coastline (Fig. 2). Both models suggest a disparate distribution of cnidarians but this is similar to that found in many modern members of this group. Therefore, the most convincing palaeobiogeographic reconstructions of the late Precambrian land mass suggest a faunal province locus²⁶ at about 30°S but the faunal distribution in the two models²⁻⁴ is equally plausible.

I thank Drs J. D. A. Piper, C. R. C. Paul and M. D. Hendry for critically reading an early draft of this paper, which was also improved following constructive comments by Dr L. Garwin.

- Piper, J. D. A. in *The Caledonide Orogen-Scandinavia and Related Areas* (eds Gee, D. G. & Sturt, B. A.) 19-34 (Wiley, Chichester, 1985).
- Piper, J. D. A. *Palaeomagnetism and the Continental Crust* 384 (Wiley, New York, 1987).
- Bond, G. C., Nickeson, P. A. & Kominz, M. A. *Earth planet. Sci. Lett.* **70**, 325-345 (1984).
- Smith, A. G., Hurley, A. M. & Briden, J. C. *Phanerozoic Palaeocontinental World Maps* 102 (Cambridge University Press, 1981).
- Hughes, N. F. (ed.) *Spec. Pap. Palaeont.* **12**, 334 (1973).
- Whittington, H. B. *Spec. Pap. Palaeont.* **12**, 235-240 (1973).
- Williams, A. *Spec. Pap. Palaeont.* **12**, 241-269 (1973).
- Glaessner, M. F. in *Treatise on Invertebrate Paleontology, Part A* (eds Robison, R. A. & Teichert, C.) A79-A118 (University of Kansas Press, 1979).
- Cope, J. C. W. *Nature* **268**, 624 (1977).
- Gibson, G. G., Teeter, S. A. & Fedonkin, M. A. *Geology* **12**, 387-390 (1984).
- Hofmann, H. J., Mountjoy, E. W. & Teitz, M. W. *Geology* **13**, 819-821 (1985).
- Dyson, I. A. *Nature* **318**, 283-285 (1985).
- Narbonne, G. & Hofmann, H. J. *4th N. Am. Paleont. Conv. Abstr.* A32-A33 (1986).
- Weiguo, S. *Precamb. Res.* **31**, 325-360 (1986).
- Anderson, M. M. & Misra, S. B. *Nature* **220**, 680-681 (1968).
- Misra, S. B. *Geol. Soc. Am. Bull.* **80**, 2133-2140 (1969).
- Cocks, L. R. M. & Fortey, R. A. *J. geol. Soc. Lond.* **139**, 465-478 (1982).
- Mayer, A. G. *Medusae of the world, III, The Scyphomedusae* 499-735 (1910).
- Campbell, A. C. *Guide to the Seashore and Shallow Seas of Britain and Europe* 320 (Country Life, London, 1976).
- Piper, J. D. A. *Phil. Trans. R. Soc. A280*, 469-490 (1976).
- Ford, T. D. *Proc. geol. Ass.* **91**, 81-83 (1980).
- Whittington, H. B. *J. Paleont.* **40**, 696-737 (1966).
- Eisbacher, G. H. *Palaeogeogr. Palaeoclimat. Palaeoecol.* **51**, 231-254 (1985).
- McMenamin, M. A. S. *Geology* **10**, 290-292 (1982).

Palaeomagnetic constraints on the collision and rotation of North and South China

Xixi Zhao & Robert S. Coe

Earth Sciences Board, University of California, Santa Cruz, California 95064, USA

The eastern half of China is dominated by the North China Block (NCB) and South China Block (SCB), separated by a complex fold system that trends roughly east-west. Interpretation of the geology of this fold belt has led to a variety of estimates of the age of amalgamation of these two blocks¹⁻⁷. Palaeomagnetic results are few and have also been subject to conflicting interpretations⁸⁻¹¹. Here we present our own palaeomagnetic results for late Permian rocks for both the SCB and NCB and summarize other relevant palaeomagnetic data. By reinterpreting the polarity given in some previous studies we are able to synthesize the results for the widespread Emeishan basalts of the SCB and to suggest a new model for the collision and relative rotation of the two blocks.

Most palaeomagnetic data of Permian age for the SCB are derived from the Emeishan Basalt Formation, which is widespread throughout south-west China. In 1980 one of us (X.Z.) sampled 32 lava flows spanning 212 m of section at the type locality (29.6°N, 103.4°E) on the east limb of the Nubeishan anticline near Emei Mountain, Sichuan Province. Table 1 shows the results of this study¹². Both alternating-field and thermal stepwise demagnetization experiments were conducted on these basalts. In general, both types of treatment were equally successful, and a stable component of magnetization was isolated that decayed univectorially towards zero. The directions of this stable component are clustered about a mean with northeasterly declination and shallow upward inclination. Here we interpret these data as primary directions acquired in a field of normal polarity.

The results of four other studies^{9,10,13,14} of the Emeishan basalts in the same area have revealed very similar north-east and upward palaeomagnetic directions (Fig. 1). The data for one of these studies⁹ pass the fold test, which means that the magnetization pre-dates the folding, a necessary (but not sufficient) requirement for the directions to be primary. In all four studies the polarity assigned is opposite to our interpretation; that is, reversed. The reason for the reversed polarity choice is that the age of the Emeishan basalt was thought to lie within

Received 11 September 1986; accepted 11 February 1987.

- Glaessner, M. F. *The Dawn of Animal Life, A Biohistorical Study*, 244 (Cambridge University Press, 1984).
- Morel, P. & Irving, E. J. *Geol.* **86**, 535-561 (1978).

Table 1 Summary of palaeomagnetic results on rocks from the SCB and the NCB

Rock type and locality	Site lat.	Site long.	<i>N</i>	Dec.	Inc.	<i>a</i> ₉₅	<i>k</i>	Pole lat.	Pole long.	<i>A</i> ₉₅
South China Block										
Emeishan Basalt, Emei, Sichuan (ref. 12)	29.6°	103.4°	32†	20.7°	−12.6°	4.0°	38.6	49.0°	251.0°	2.1°
North China Block										
Redbeds, Jingle, Shanxi	38.6°	112.1°	11	316.1°	31.5°	10.0°*	45.6	46.2°	5.2°	10.1°*
Sandstone and siltstone, Kouquan, Shanxi	40.1°	113.2°	13	347.9°	35.5°	15.1°*	20.6	67.1°	323.6°	14.0°*
Sandstone, Lincheng, Hebei	37.5°	114.4°	17	324.7°	34.4°	11.3°*	36.4	54.3°	4.1°	8.7°*

N, number of samples used; Dec., declination; Inc., inclination; *a*₉₅: semiangle of the cone of 95% confidence; *k*, precision parameter for directions; *A*₉₅, circle of 95% confidence limit of pole.

* Confidence limits obtained by ordinary Fisher statistics from 50% of the samples that exhibited stable endpoints.

† Number of flows sampled.

the late Palaeozoic Kiaman Reversed Interval¹⁵. Assignment of reversed polarity to these palaeomagnetic directions with north-east declinations, however, implies that the area has undergone an azimuthal rotation since Permian time of about 160°. McElhinny *et al.*⁹ originally interpreted this as a rotation of the entire SCB. Chan *et al.*¹⁰, however, pointed out that the area sampled might have undergone a local rotation because it lies within the Lungmenshan thrust zone.

Recent studies of the Emeishan basalts in other regions reveal that some flows have directions with southwesterly declinations and shallow downward inclinations. Lin *et al.*¹¹ reported such a direction from a single lava flow in Guizhou Province. They interpreted this magnetization as reversed, attributing the apparently conflicting earlier result⁹ from the Emei Mountain region to secondary remagnetization or local rotation. Later studies^{16,17} of the Emeishan Basalt from southern Sichuan and Yunnan Provinces report directions for sequences of flows with both northeasterly and southwesterly declinations. Particularly diagnostic is the latter, in which the upper flows have northeasterly and upward directions and the lowest flow has a southwesterly and downward direction. The underlying, Lower Permian bauxite beds also have southwesterly declinations. The simplest interpretation is that the Emeishan Basalt Formation does not lie exclusively within the Kiaman Reversed Interval, and that the results record both normal (northeasterly declination) and reversed (southwesterly declination) polarity of the ancient field. In this way we accept the results of all studies and eliminate the need for exceptionally large azimuthal rotations. This interpretation is also one of three possible explanations of Emeishan data offered by Huang *et al.*¹⁶.

Recently published radiometric dates from the Emeishan Basalt support this hypothesis. Yuan *et al.*¹⁸ find an age range of 230–280 Myr for the basalt, with a tendency toward older ages to the south. They cite a mean age of 236 Myr for flows from Sichuan Province, which corresponds to the upper limit of the Kiaman Reversed Interval^{19,20}. If these ages are accurate, it is not surprising that both normal and reversed polarity are found and that, in particular, the polarity recorded by the flows at the type locality near Emei Mountain is normal.

Figure 1 shows results for the various studies of the Emeishan Basalt. In general these directions agree very well with each other and, as discussed above, some of them pass fold and reversal tests. For these reasons, and because of the wide distribution of sampling sites (see Fig. 3a), we believe that these are primary directions suitable for calculating a late-Permian pole position for the SCB. We have done this in three different ways: (1) averaging by regions (*N* = 5), (2) averaging by studies (all studies, *N* = 8), and (3) averaging by selected studies (omitting two results with very small numbers of samples^{10,11}, *N* = 6). All three methods yield similar results, but we prefer the third

because the mean directions used to obtain the poles are more nearly comparable in terms of numbers of samples and flows.

Next, we sampled sedimentary sections of the Shihezi Formation at three widely separated areas in Shanxi and Hebei provinces in the NCB. In Shanxi province more than 800 m of Permian sandstone and siltstone are exposed in the Kouquan area (40.1° N, 113.2° E) of the Datong Basin. In the Jingle area (38.6° N, 112.1° E) the same formation is well exposed on both limbs of a syncline. In the Lincheng area (37.5° N, 114.4° E) of Hebei province, fresh exposures of a sandstone unit crop out along the banks of the Shaba ravine. At all three localities the coal-bearing Shanxi Formation, which is early Permian in age, can be easily recognized at the base of the sections. The strata we sampled are palaeontologically dated as late Permian²¹, and are mainly yellowish except at Jingle, where redbeds are common. Rocks from all three areas exhibited similar magnetic behaviour in the laboratory. Progressive thermal demagnetization of the samples from each locality revealed, in ~50% of the cases, a characteristic component of magnetization with southeasterly or northwesterly declinations and shallow upward or downward inclinations. We also found evidence for this component in the demagnetization planes of the remaining samples. Reanalysis of our data using the joint demagnetization line and plane method of Kirschvink²² yielded a direction that is not distinguishable at 95% confidence from the characteristic component.

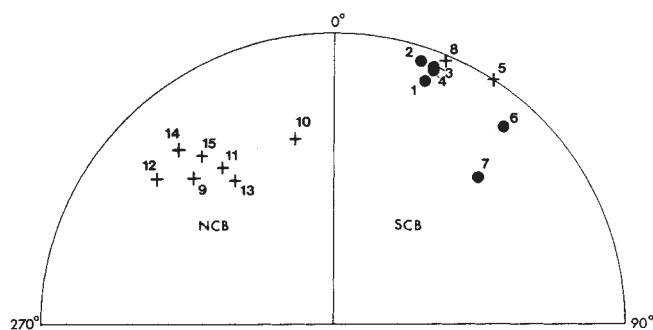


Fig. 1 Site mean directions of Permian rocks from the SCB and NCB (equal-area projection with circles and crosses representing upper and lower hemisphere, respectively, and with reversed directions inverted to normal). On the SCB all sites are Emeishan Basalt: (1) ref. 12 (Table 1); (2) ref. 9; (3) ref. 14; (4) ref. 13; (5) ref. 16; (6) ref. 17; (7) ref. 11; (8) ref. 10. On the NCB, sites are in redbeds, sandstone, and siltstone: (9) Jingle redbeds (Table 1); (10) Kouquan sandstone and siltstone (Table 1); (11) Lincheng sandstone (Table 1); (12) ref. 11; (13) W. Z. Lin, J. Shao and Z. Zhao, unpublished data; (14) ref. 9; (15) ref. 13.

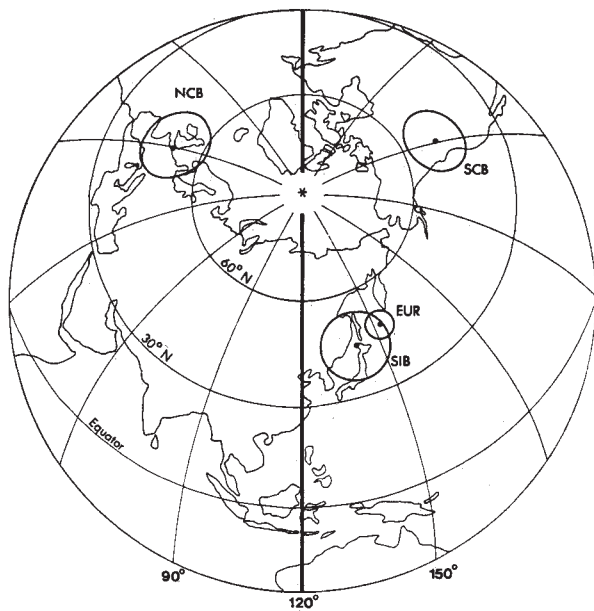


Fig. 2 Permian pole positions. NCB: 51.4° N, 359.8° E, $A_{95} = 9.3$; SCB: 49.8° N, 240.5° E, $A_{95} = 8.4$; Siberia (SIB): 45.0° N, 141.0° E, $A_{95} = 10.0$ (ref. 29); and Europe (EUR): 48.0° N, 153.0° E, $A_{95} = 4.0$ (ref. 30).

Table 1 gives the palaeomagnetic results of this analysis. A detailed account of these studies will be published elsewhere. The site mean directions pass the fold test at $>95\%$ confidence. Thus the remanent magnetization at these localities was acquired before folding, which occurred in the middle Jurassic²¹. Moreover, the normal and reversed directions that we find are closely antiparallel. The successful fold and reversal tests, and the fact that this characteristic component is carried by both magnetite and haematite, strongly suggest that the magnetization is primary.

Four other Permian palaeomagnetic results are available for the NCB (refs 9, 11, 13 and W. Z. Lin, J. Shao and Z. Zhao, unpublished data). These are all derived from a redbed sequence of late Permian age near Taiyuan, Shanxi Province. They agree reasonably well among themselves and with our results from rocks to the east and west (Fig. 1).

The Permian palaeomagnetic directions for the SCB and NCB shown in Fig. 1 are completely different from each other. The corresponding mean palaeomagnetic poles computed from these directions (Fig. 2) are distinct at $>99\%$ confidence. Because the data for both the SCB and NCB pole positions pass fold and reversal tests and are mutually consistent, we believe they represent primary remanence with little or no contamination by unremoved secondary components. Thus we must look for tectonic causes to explain the discrepancy. Moreover, these directions are representative of large areas on the SCB and NCB (Fig. 3a), so that local rotation of small blocks is not a likely explanation.

The palaeolatitudes implied by the data from the SCB and NCB are similar, but the mean direction of palaeo-north for the SCB is rotated clockwise more than 60° with respect to that for the NCB. Thus the NCB and SCB could not have been in their present configuration in late Permian time. A model involving relative rotation of the SCB and NCB is suggested by the geometrical relationships evident in Fig. 2. The great circle equidistant between the SCB and NCB poles intersects the boundary between these two blocks near its eastern end. A finite rotation of 67° about this point of intersection brings the SCB and NCB poles into coincidence. As Fig. 3 shows, this is what might be expected if collision occurred initially at this point in easternmost China and progressed westward as the SCB rotated

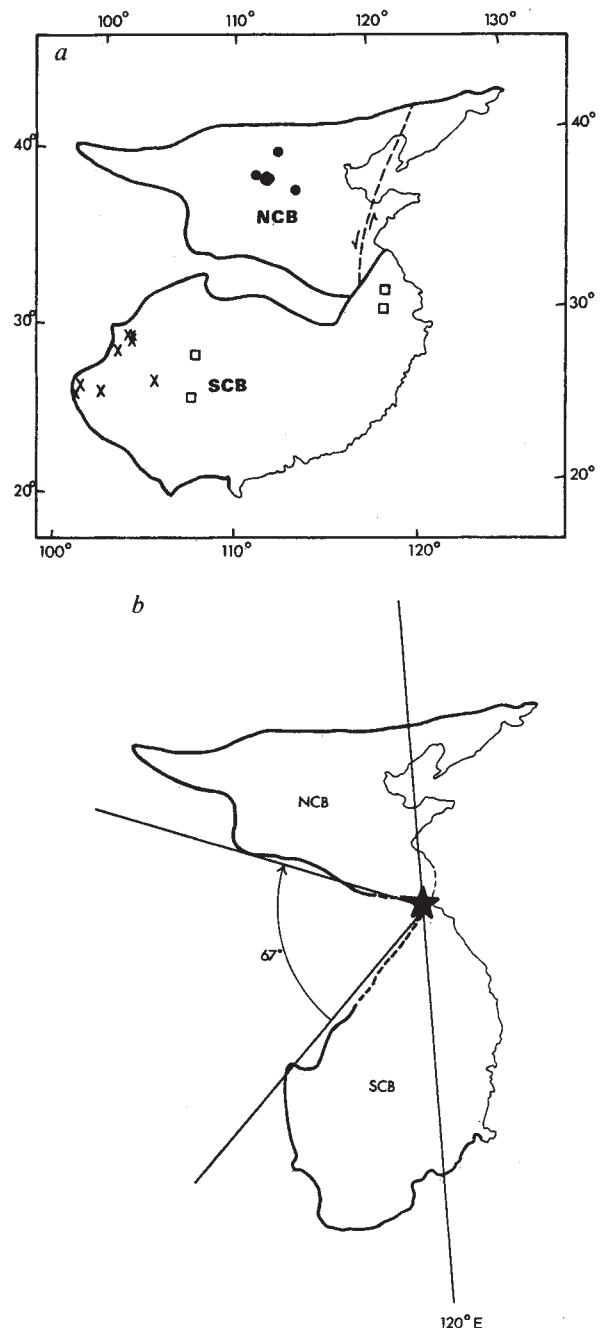


Fig. 3 a, Present location of the North and South China blocks. Thin-line boundaries are shorelines (not block boundaries) and the broken line is the Tancheng-Lujiang fault. The sampling sites from the NCB and SCB are also indicated with squares representing Triassic sites, and circles and crosses representing Permian sites. b, Schematic reconstruction of the NCB and SCB in Permo-Triassic time. The NCB is arbitrarily kept in its present orientation relative to north. The boundaries are simplified in the collision zone, and the Tancheng-Lujiang Fault is omitted.

clockwise relative to the NCB. The distribution of Triassic flysch deposits suggests that this model merits serious consideration. These are well developed along the northwestern margin of the SCB, whereas on the northeastern margin there are only littoral and subaerial deposits of that age²⁴.

Of course, a different finite rotation about any other point on the bisecting great circle (Fig. 2), or any number of more complex combinations of rotations, can bring the palaeomagnetic poles for the SCB and NCB into coincidence. An interesting case is rotation about a point chosen so that the relative displacement

between the SCB and NCB along their boundary is approximately strike-slip. The amount of displacement needed to accomplish the azimuthal rotation required by the palaeomagnetic data, however, is very large: more than 5,000 km. We know of no geological evidence for such displacement.

Early Triassic data for the NCB and SCB are similar to those for the late Permian. Two studies^{11,13,23} near Taiyuan in Shanxi Province yield directions that agree well with those from the late Permian rocks from the same area. A third result^{11,23} from the Benxi area, far to the east in Liaoning Province, does not agree with the other two for the NCB, even after correction of a mistake in the calculation of the mean direction and pole. The inclination is concordant with the other two, but the declination is rotated about 40° anticlockwise, in agreement with the sense of rotation that would be expected from the nearby Tancheng-Lujiang Fault. The possibility of such local rotations near this major fault zone has been suggested by others^{25,26}. For this reason, and because no evidence was presented demonstrating that the magnetization is primary, we exclude this result from further consideration. On the SCB, one middle Triassic and three early Triassic results^{10,11,27} agree well with each other and are similar to the late Permian results. Because of their wide distribution (Fig. 3a), they further support the contention that the palaeomagnetic directions are representative of the whole SCB.

Thus the early Triassic data may be explained in the same way as those of the late Permian. If we average the two sets of data to obtain Permo-Triassic poles for the SCB and NCB, the model illustrated in Fig. 3 applies with only minor modification. The main change is that the point of initial collision about which the blocks rotate is shifted several hundred kilometres westward along the SCB-NCB boundary.

The question of when the SCB and NCB collided and became completely sutured still cannot be answered precisely. Well-documented studies show that the SCB and NCB palaeomagnetic poles are statistically indistinguishable in late Cretaceous time²⁸. Therefore, collision and suturing was complete before then. Middle and late Jurassic pole positions^{11,23} have been published that are similar for both blocks, but no documentation was offered to demonstrate that these rocks escaped Cretaceous overprinting that is so pervasive in eastern China. We know of no data for the late Triassic and early Jurassic. Geological evidence is even less definitive, with estimates ranging from Devonian to Jurassic¹⁻⁷. We favour the hypothesis that North and South China initially collided in the early Triassic and finished most of their relative rotation during the Jurassic.

A final question concerns the time of amalgamation of China and Siberia. Some geological^{2,4} and palaeomagnetic¹¹ articles have suggested that the NCB accreted to Siberia before the late Permian. The late Permian (Fig. 2) and early Triassic palaeomagnetic data contradict these interpretations. As pointed out first by McElhinny *et al.*⁹ in 1981 and most recently by Opdyke *et al.*²⁷ in 1986, these data imply that both the SCB and the NCB lay far to the south compared with their present positions relative to Siberia. Where and when the relative displacement between China and Siberia was accommodated remains an important problem in the tectonics of Asia.

We thank the Geological Bureau of Shanxi and Sichuan, the Institute of Geology of the Academia Sinica, the Institute of Geology of the National Seismological Bureau, and the Seismological Bureau of Hebei Province, for their help in our study; Zhou Yaoxiu, Ye Sujuan, Wang Jie, Han Jiamao, Deng Yuren and Zhang Huanqun for help with the field work; and Zhou Yaoxiu, Liu Chun and Dennis Kent for details of their palaeomagnetic results. This study was supported by NSF grant EAR-8120368.

3. Mattauer, M. *et al.* *Nature* **327**, 496-500 (1985).
4. Zhang, W. Y. & Zhang, C. *Research on Geology* Vol. 1, 10-21 (Inst. Geol. Academia Sinica, Cultural Relics Publishing House, 1982).
5. Li, C. Y., Wang, Q., Liu, X. Y. & Tang, Y. Q. *Bull. Chin. Acad. geol. Sci.* No. 10, 3-11 (1984).
6. Wang, H. Z., Xu, C. Y. & Zhou, Z. G. *Acta geol. sin.* No. 3, 270-280 (1982).
7. Yang, S. N. & Yang, W. L. *Regional Geotectonics of China* (Geological Publishing House, Beijing, 1985).
8. Wang, T., Teng, H., Li, C. & Ye, S. *Acta geophys. sin.* **9**, 125-138 (1960).
9. McElhinny, M. W., Embleton, B. J. J., Ma, X. H. & Zhang, Z. K. *Nature* **293**, 212-216 (1981).
10. Chan, L. S., Wang, C. Y. & Wu, X. Y. *Geophys. Res. Lett.* **11**, 1157-1160 (1984).
11. Lin, J. L., Fuller, M. & Zhang, W. Y. *Nature* **313**, 444-449 (1985).
12. Zhao, X. thesis, Inst. Geol. Academia Sinica, Beijing (1981).
13. Zhang, Z. K. *Bull. Chin. Acad. geol. Sci.* No. 4, 45-54 (1984).
14. Liu, C., Zhu, R. & Jing, Z. *Monograph of China Panxi Rift* Vol. 2 (Geological Publishing House, Beijing) (in the press).
15. Irving, E. & Parry, L. G. *Geophys. J. R. astr. Soc.* **9**, 153-167 (1963).
16. Huang, K., Opdyke, N. D., Kent, D. V., Xu, G. & Tang, L. *Kexue Tongbao* **31**, 1195-1201 (1986).
17. Zhou, Y. X. *et al.* *Dizhi Pinglun* **32**, 465-469 (1986).
18. Yuan, H. *et al.* *Monograph of China Panxi Rift* Vol. 1 (Geological Publishing House, Beijing, 1985).
19. Harland, W. B. *et al.* *A Geological Time Scale* (Cambridge University Press, 1982).
20. Khramov, A. N. & Rodionov, V. P. in *Global Reconstruction and the Geomagnetic Field during the Palaeozoic* (eds McElhinny, M. W. *et al.*) (*Adv. Earth planet. Sci.* **10**) 99-115 (1981).
21. Chinese Geological Academy *Geological Atlas of People's Republic of China* (1973).
22. Kirschvink, J. *Geophys. J. R. astr. Soc.* **62**, 699-718 (1980).
23. Lin, J. L. thesis, Univ. California Santa Barbara (1984).
24. Yang, Z. Y. *et al.* *An Outline of the Stratigraphy of China* (ed. Chinese Academy of Geological Sciences) (Geological Publishing House, Beijing, 1982).
25. Courtillot, V. & Besse, J. *Nature* **320**, 86-87 (1986).
26. Kent, D. V., Zeng, X. S., Zhang, W. Y. & Opdyke, N. D. *Tectonophysics* (in the press).
27. Opdyke, N. D., Huang, K., Xu, G., Zhang, W. Y. & Kent, D. V. *J. geophys. Res.* **91**(B9), 9553-9568 (1986).
28. Kent, D. V., Xu, G., Huang, K., Zhang, W. Y. & Opdyke, N. D. *Earth planet. Sci. Lett.* **79**, 179-184 (1986).
29. Khramov, A. N. *Paleomagnetism*, 303 (Leningrad, 1982) (in Russian).
30. Irving, E. & Irving, G. A. *Geophys. Surv.* **5**, 141-188 (1982).

Natural enemies may be a cause of discrete generations in tropical insects

H. C. J. Godfray* & M. P. Hassell†

Department of Zoology, University of Oxford, Oxford OX1 3PS, UK
† Department of Pure and Applied Biology, Imperial College at Silwood Park, Ascot, Berkshire SL5 7PY, UK

It is commonly believed that the marked seasonality of temperate regions results in the insects that live there often having discrete generations, whereas in the tropics, generations are more likely to overlap because of the more-or-less continuously favourable conditions. These sweeping generalizations, however, are misleading. Several examples are now known of tropical insect populations that tend to show discrete generations, both under equilibrium¹⁻⁵ (Fig. 1a) and non-equilibrium conditions⁶⁻⁹, which are not seasonally determined. We are also struck by the frequency with which laboratory insect populations cultured under constant conditions show discrete generations¹⁰⁻¹² (Fig. 1b). A common factor in these examples is the presence of insect parasitoids as a major mortality factor. Here we demonstrate, using theoretical population models, one way in which parasitism can be the cause of approximately discrete generations. In particular, we show that the ratio of the lengths of the host and parasitoid life cycles is of prime importance in determining whether generations tend to be discrete or continuous and discuss the importance of this in biological control.

A popular approach to the modelling of host-parasitoid interactions has been the use of non-linear difference equation models¹³⁻¹⁶. These coupled, discrete-generation models are particularly appropriate to temperate regions where seasonality often synchronizes the populations and provides a natural interval between the appearance of successive generations. Such models, however, are not well suited to tropical host-parasitoid

Received 14 October 1986; accepted 5 January 1987.

1. Klimetz, M. P. *Tectonics* **2**, 139-166 (1983).
2. Zhang, Z. M., Liou, J. G. & Coleman, R. G. *Geol. Soc. Am. Bull.* **95**, 295-312 (1984).

* Present address: Department of Pure and Applied Biology, Imperial College, Silwood Park, Ascot, Berks, UK.

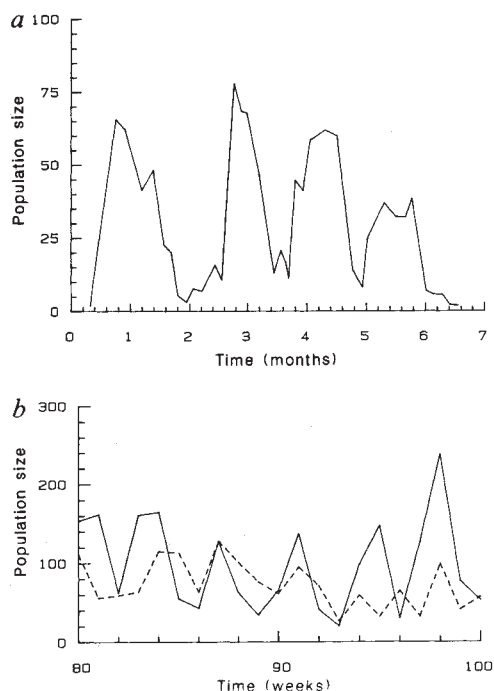


Fig. 1 Examples of approximately discrete generations in insect populations. *a*, Part of a time series of emergences of the adult coffee leaf miner, *Leucoptera coffeina* Wshbn, from field samples of pupae⁴. *b*, Laboratory populations of the bruchid beetle, *Callosobruchus chinensis* (L.) (solid line) and the pteromalid parasitoid, *Anisopteromalus calandrae* Howard (broken line), taken from weeks 80 to 100 of a long-term interaction²⁵. In both the field^{26,27} and the laboratory²⁵ system the parasitoids have generation times of the order of one half that of their host.

systems where the generation times of hosts and parasitoids are frequently quite different. Such situations call for models in which the age-structures of host and parasitoid are explicitly included. A classic study of this kind, using models framed as delay-differential equations, is that of Auslander *et al.*¹⁷ who found that discrete generation 'waves' could occur in an age-structured, coupled host-parasitoid model depending sensitively on several parameters (for example, the extent of refuges from parasitism available to the hosts and the preference of parasitoids for different host stages). More recently, Murdoch *et al.*¹⁸

obtained discrete generation behaviour in an age-structured model of the population dynamics of the red scale and its parasitoids.

Our approach has been to start with a minimally complicated host population model in which the lengths of the juvenile pre-parasitism, susceptible and post-parasitism stages, and of the adult stage are specified, and where adults have a constant fecundity per unit time. We assume that: (1) the length of the pre- and post-parasitism stages are either constant or drawn from an inverse quadratic distribution with specified mean and variance¹⁹; (2) the duration of the susceptible stage is fixed at one time unit; and (3) the adult longevity is either also fixed or there is a constant probability of death giving an exponential (type II) survivorship curve. The type of insect we have in mind therefore resembles a typical tropical moth in which the length of the susceptible and adult stages are only a small part of the total generation time. Our host population is clearly unlimited and, as long as cohorts mix, will always move towards completely overlapping generations.

To this picture we now add a monophagous parasitoid population whose life cycle is divided into two parts: the adult reproductive stage that parasitizes host larvae and the immature stages developing on or in parasitized hosts. The lengths of the adult and juvenile stages are determined in a similar way as those of the host. We assume the probability of a host larva surviving parasitism to be given by the zero term of the negative binomial distribution²⁰⁻²⁴:

$$f = (1 + aP_t/k)^{-k} \quad (1)$$

Here a is the per capita searching efficiency of the parasitoids, P_t is the number of adult parasitoids searching during time t and k is a measure of the degree of contagion in the distribution of parasitism among hosts. Equation (1) provides a simple means of exploring some of the effects of non-random parasitism in population models²¹.

One limiting case of this model occurs when the total host and parasitoid life cycles are of the same length, the adult hosts and parasitoids live for only one time unit, and there is no variance in the non-reproductive stages. The model now becomes:

$$\begin{aligned} N_{t+1} &= FN_t f(P_t) \\ P_{t+1} &= N_t [1 - f(P_t)] \end{aligned} \quad (2)$$

where F is the per capita reproductive rate per generation and f is given by equation (1). The properties of equation (2) are

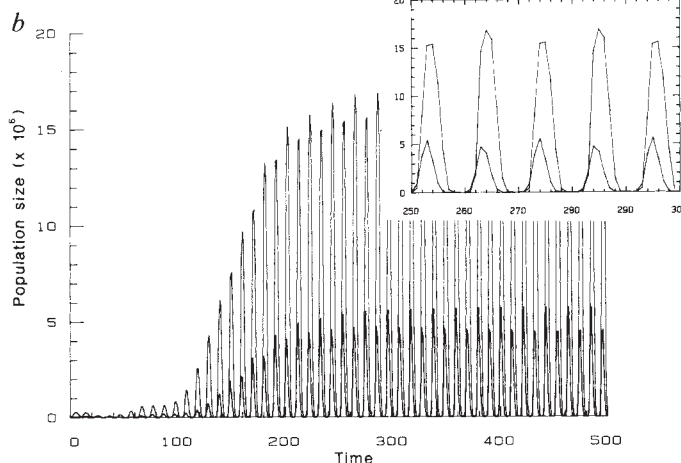
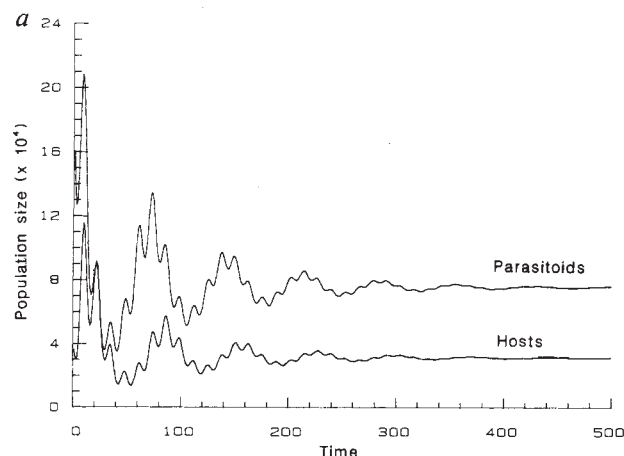


Fig. 2 Examples of 'continuous' and 'discrete' generation behaviour given by the model. *a*, Hosts and parasitoids with equal life-cycle lengths; *b*, parasitoid life cycle is less than that of the host (inset, detail of generation cycle over 50 time units; axes as in main graph). Parameters common to both plots are daily host adult fecundity ($F=8$), parasitoid searching efficiency ($a=0.001$) and contagion in the distribution of parasitism among hosts ($k=0.8$). In both plots the duration of the host immature stages are: pre-parasitism, 4 time units; susceptible, 1 time unit; post-parasitism, 4 time units (all with zero variance). Adult host and parasitoid longevity are exponentially distributed with mean 2 time units except that all remaining individuals are assumed to die when the probability of survival falls to 0.1. In *a*, the juvenile parasitoid stage is 9 time units in duration whereas in *b* it is 5 time units.

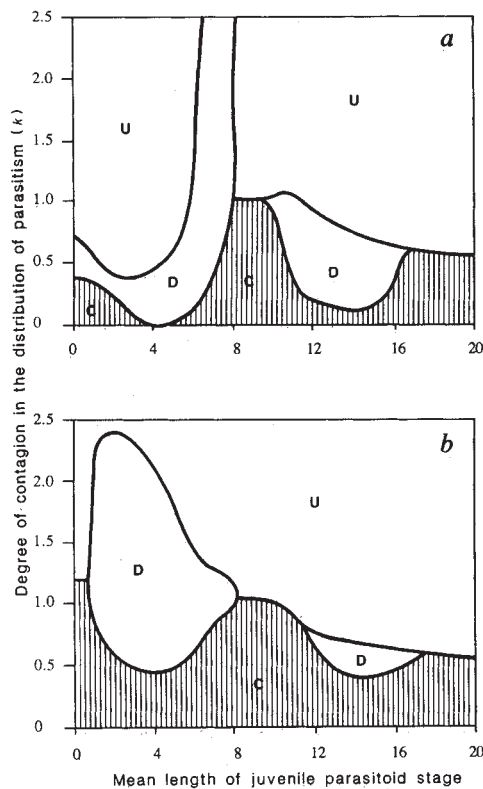


Fig. 3 Population behaviour as a function of the degree of contagion in the distribution of parasitism among hosts (k) and the difference in length of host and parasitoid life cycles. *a*, Cohort overlap occurs through distributed development periods. The lengths of the four host stages, with their variance in parentheses, are 4(2), 1(0), 4(0), 1(0) time units respectively. The parasitoid adult lives for one time unit and the mean length of the juvenile parasitoid stage is given on the x-axis with variance half the mean. *b*, Cohort overlap occurs because adults live for longer than one time unit. The lengths of host and parasitoid life cycles are the same as in Fig. 2 with the juvenile parasitoid stage varying on the abscissa. Three types of population behaviour are distinguished: U, unstable; D, discrete generations at equilibrium; C, (hatched area), continuous generations at equilibrium. In both *a* and *b*, $F = 8$ and $a = 0.001$.

straightforward: there is a stable equilibrium for all $k < 1$ (that is, if there is sufficient contagion in the distribution of parasitoid attacks among hosts)²¹. This result remains true in numerical simulations if, keeping the lengths of the host and parasitoid life cycles the same, overlap of cohorts is allowed either by introducing distributed developmental periods of the immature stages, or by allowing adults to live for more than a single time unit. The only difference from equation (2) is that the populations reach an equilibrium with all host and parasitoid stages simultaneously present (complete overlap of generations).

To explore the effects of unequal host and parasitoid life-cycle lengths, we keep that of the host constant and manipulate the length of the parasitoid life cycle by altering the duration of its immature stage. In numerical simulations, two distinct types of population behaviour are observed at equilibrium: (1) constant population sizes with all host and parasitoid age classes present—continuous generations (Fig. 2*a*); and (2) stable cycles in both host and parasitoid populations with periods of both species approximately equal to one host generation—discrete generations (Fig. 2*b*). The two most important parameters in determining which of these behaviours occurs is the ratio of life-cycle lengths and the degree of aggregation (k of the negative binomial), as shown in Fig. 3. Very low values of k always promote continuous generations at equilibrium, but when aggregation is somewhat less strong this is only observed when the

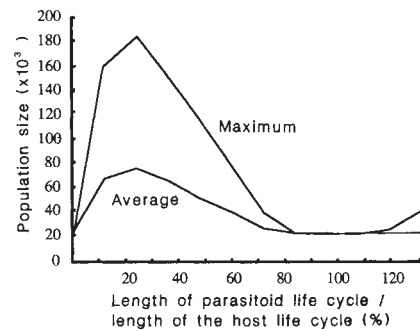


Fig. 4 Equilibrium host density as a function of the difference between host and parasitoid life cycles. Two measures of host density are used: the average host density at equilibrium (lower curve) and the maximum host density (upper curve). These are identical if the generations are continuous at equilibrium. Parameter values as in Fig. 2 and Fig. 3*b*. The relative lengths of host and parasitoid life cycles are altered by changing the length of the juvenile parasitoid stage.

parasitoid life cycle is very short, about the same length as the host or about twice that of the host. Discrete generations, on the other hand, are most likely when the life cycle of the parasitoid is about half (Fig. 1), or 1.5 times that of the host. This pattern is found whether cohort overlap occurs through distributed immature development periods or by allowing adults to live for more than a single time unit. Changes in other parameters affect the relative sizes of the regions in Fig. 3 without affecting the overall pattern. Reducing adult host fecundity, for example, contracts the region of discrete generations. Further effects of varying parasitoid generation times are on the region of stability (Fig. 3) and on equilibrium population levels (Fig. 4), with higher levels when populations are discrete.

A verbal explanation for the mechanism underlying the patterns in Fig. 3 is as follows. At any time, a large host population will tend to result in the host population again being high one host generation interval later. This will also result in increased numbers of parasitoids one parasitoid generation later. If the two generation times are more-or-less the same, this will lead to dampening of the host generation cycles and a move towards constant host and parasitoid populations at equilibrium. A similar argument holds for parasitoids with a generation time twice that of the host. However, if the parasitoid has a generation time half that of the host, the increased numbers of parasitoids stemming from a peak host population will emerge half a host generation later, to inflict a heavy host mortality and accentuate the trough in host densities. Discrete generations are thus promoted. The same mechanism would apply if the parasitoid generation time is 1.5 times that of the host. Should the parasitoid generation time be very small, the numerical response will be almost immediate, rather than delayed, which provides the continuous feedback necessary to dampen the generation cycles of the host. Finally, when the distribution of parasitism among hosts is very clumped ($k \ll 1$ in equation (1)), a large proportion of the host population will be unaffected by parasitism and the disruptive effect of parasitoid and host peaks not coinciding is much smaller. The presence of any such refuge, implicit or explicit, is likely to have much the same effect of promoting continuous populations.

The effect of different generation times of hosts and parasitoids described above has potentially important implications for biological control where a chosen parasitoid is released to control a pest population. There may be an advantage to continuous generations stemming from the lower equilibrium population levels (Fig. 4). On the other hand there could be a net advantage in having discrete generations. If supplementary chemical pest control is used, it may be beneficial to have host and parasitoid populations synchronized for optimal timing of

chemical application. In addition there will be differences in the behaviour of the two types of population following natural or artificial perturbation from equilibrium. Discrete generation populations are more susceptible to perturbation and have a longer return time. In the past, characterizations of an 'ideal' biological control agent predicted from host-parasitoid population models have concentrated on the searching performance of the parasitoid which is notoriously difficult to estimate in the field. We suggest that the relative lengths of host and parasitoid generations also have a profound impact on population dynamics and yet are one of the easiest attributes to measure under natural conditions.

We thank John Lawton and Jeff Waage for helpful comments. We are grateful for grants from the Council of the European Community and thank the Natural Environmental Research Council for a grant to M.P.H.

Received 14 January; accepted 11 March 1987.

1. Bigger, M. *Nature* **259**, 207-209 (1976).
2. Taylor, T. H. C. *The Biological Control of an Insect in Fiji* (Imperial Institute of Entomology, London, 1937).
3. Metcalfe, J. R. *Bull. ent. Res.* **60**, 565-597 (1971).

4. Notley, F. B. *Bull. ent. Res.* **46**, 899-899 (1955).
5. Banerjee, B. *Bull. ent. Res.* **69**, 195-201 (1979).
6. Tothill, J. D., Taylor, T. H. C. & Paine, R. W. *The Coconut Moth in Fiji* (Imperial Institute of Entomology, London, 1930).
7. Van der Vecht, J. *Ent. Ber. Amst.* **15**, 122-132 (1954).
8. Wood, B. J. *Pests of Oil Palms in Malaysia and Their Control* (Incorporated Society of Planters, Kuala Lumpur, 1968).
9. Perera, P. A. C. R. thesis, Univ. London (1987).
10. Utida, S. *Ecology* **38**, 442-449 (1957).
11. Hassell, M. P. & Huffaker, C. B. *Res. Popn Biol.* **11**, 186-210 (1969).
12. White, E. G. & Huffaker, C. B. *Res. Popn Biol.* **11**, 150-185 (1969).
13. Hassell, M. P. *The Dynamics of Arthropod Predator-Prey Systems* (Princeton University Press, 1978).
14. Wang, Y. H. & Gutierrez, A. P. *J. Anim. Ecol.* **49**, 435-452 (1980).
15. May, R. M., Hassell, M. P., Anderson, R. M. & Tonkyn, D. W. *J. Anim. Ecol.* **50**, 855-865 (1981).
16. Hassell, M. P. & May, R. M. in *Behavioural Ecology, Ecological Consequences of Adaptive Behaviour* (eds Sibley, R. M. & Smith, R. H.) *Brit. Ecol. Soc. Symp.* **25**, 3-32 (1985).
17. Auslander, D. M., Oster, G. F. & Huffaker, C. B. *J. Franklin Inst.* **297**, 345-376 (1974).
18. Murdoch, W. W., Nisbet, R. M., Blythe, S. P., Gurney, W. S. C. & Reeve, J. D. *Am. Nat.* (in the press).
19. Sharpe, P. J. H., Curry, G. L., De Michele, D. W. & Cole, C. L. *J. theor. Biol.* **66**, 21-38 (1977).
20. Griffiths, K. J. & Holling, C. S. *Can. Ent.* **101**, 785-818 (1969).
21. May, R. M. *J. Anim. Ecol.* **47**, 833-843 (1976).
22. Chesson, P. L. & Murdoch, W. W. *Am. Nat.* **127**, 696-715 (1986).
23. Hassell, M. P. *J. Anim. Ecol.* **49**, 603-628 (1980).
24. Perry, J. N. & Taylor, L. R. *J. Anim. Ecol.* **55**, 1053-1068 (1986).
25. May, R. M. & Hassell, M. P. *Phil. Trans. R. Soc.* (in the press).
26. Notley, F. B. *Bull. ent. Res.* **39**, 399-416 (1948).
27. Bigger, M. *J. Anim. Ecol.* **42**, 417-434 (1973).

Demographic study of a wild house sparrow population by DNA fingerprinting

Jon H. Wetton, Royston E. Carter, David T. Parkin & David Walters

Department of Genetics, Medical School, Queen's Medical Centre, Nottingham NG7 2UH, UK

Over the past twenty years, several techniques from biochemical and molecular genetics, such as enzyme electrophoresis and isoelectric focusing, have been widely and successfully applied to the study of population differentiation and evolution¹. However, they have been less applicable to demographic problems such as assigning parentage to individuals within a population. This stems from a general weakness of data derived from enzyme loci: allele frequencies at polymorphic loci are sufficiently skewed that the majority of individuals are of one or two genotypes. Many enzyme systems can only be examined post mortem, so that the loci are of little use if the animals are to be studied in the wild. The search for new and more sensitive techniques for detecting genetic variation has continued, and recently a major discovery has come from molecular biology. Jeffreys *et al.*² have reported the detection of a type of hypervariable 'minisatellite' DNA that is extraordinarily polymorphic in human populations. We have applied their technique to several bird species and particularly to a population of house sparrows (*Passer domesticus*) near Nottingham. We report here that one of the human minisatellite clones is a suitable probe for sparrow DNA and that it reveals variation as extensive as that found in man. These results suggest that analysis of minisatellite DNA will be a powerful tool in the study of demographic population genetics.

The study population of house sparrows lives around farm buildings at Brackenhurst College of Agriculture³. It has been under observation since 1979 and currently there are 80 nest boxes, of which about 60 are occupied each year. During the breeding season, the boxes are monitored at two-day intervals, and all nestlings that reach 10 days of age are ringed and blood samples are taken. Adult birds are trapped in the colony with nets, and also at the nest during breeding. Every adult bird is given a unique combination of colour rings, and a blood sample is taken if the bird has not been handled before. The use of colour rings, combined with direct observation, allows the iden-

tification of mated pairs, and of their attendance at nests during the incubation and feeding of young, without the need for continuous disturbance.

A survey of the population was undertaken using starch gel electrophoresis of seven polymorphic loci. This shows that approximately 8% of nestlings are genetically mismatched with at least one of their putative parents (ref. 3 and J.H.W., unpublished data). The allele frequencies are such that we would expect to detect about half the cases in which non-parental nest attendance occurs. However, in common with similar studies, the data are poor and errors are large. To illustrate the power of polymorphic minisatellite analysis in resolving familial relationships, we will describe the results from two separate nest boxes in the house sparrow population during 1985 and 1986.

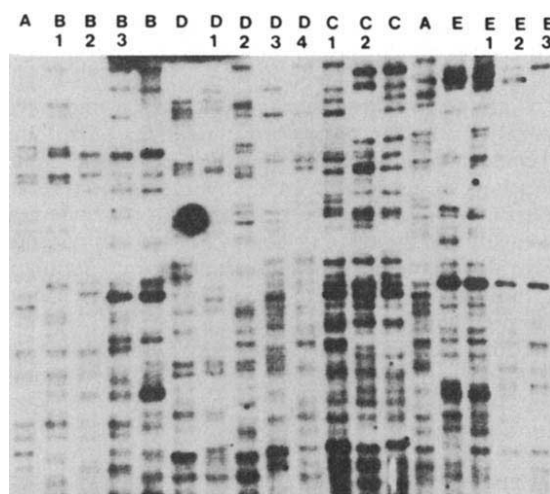
The first family consists of a male A who was paired with female B during 1985, and then with three different females (C, D and E) successively through 1986. Young birds were fledged successfully from each breeding attempt, giving rise to families of three (B1-B3), two (C1-C2), four (D1-D4) and three (E1-E3) offspring. The enzyme data give no evidence of any mismatch between parents and offspring that could be assigned to infidelity or egg-dumping (data not shown). DNA was extracted from whole blood and prepared as described in Fig. 1. After treatment with *Hae*III, electrophoresis, Southern blotting and hybridization with the 33.6 minisatellite probe from Dr Alec Jeffreys, the autoradiograph shown in Fig. 1 was produced. We can distinguish a total of about 60 bands with some confidence, of which the four unrelated females have an average of 14.25 (s.e.m. = 0.85) each. Comparing these four individuals in pairs shows that the probability that a fragment present in one is also present in another is 0.086. The probability that all fragments present in one individual are present in another is thus: $0.086^{14.25}$ ($<10^{-14}$). As with humans, house sparrows are extremely variable in their minisatellite phenotypes.

Table 1 Similarity coefficients *D* calculated from the results in Fig. 1

		Sets of offspring			
		B1-B3	C1-C2	D1-D4	E1-E3
Adults	A	0.577 (0.105)	0.453 (0.021)	0.535 (0.033)	0.615 (0.028)
	B	0.628 (0.017)	0.144 (0.085)	0.175 (0.045)	0.169 (0.036)
	C	0.077 (0.039)	0.687 (0.040)	0.156 (0.059)	0.092 (0.061)
	D	0.040 (0.020)	0.247 (0.065)	0.480 (0.064)	0.172 (0.059)
	E	0.201 (0.052)	0.066 (0.001)	0.160 (0.056)	0.506 (0.118)

Values shown are means, with s.e.m. in parentheses.

The results are more complicated, not only because there is clear evidence of incest in the family, but also because the banding patterns are more similar, even between the original parents. It is possible that there had been some inbreeding of which we are unaware earlier in the pedigree. Nevertheless, the results are striking. The progeny C-F show an average similarity



Methods. Whole blood (25 μ l) collected by jugular venipuncture was suspended in 475 μ l SET buffer (0.15 M NaCl, 0.05 M Tris, 1 mM EDTA pH 8.0) and stored at -80°C . Subsequently, 15 μ l proteinase K (10 mg ml^{-1}) and 7.5 μ l 25% w/v SDS were added to the thawed sample and it was incubated overnight at 55°C . DNA was purified by three extractions with phenol, followed by two with phenol/chloroform and one with chloroform, and was precipitated with 0.1 vols 3M sodium acetate and 2 vols absolute ethanol at 0°C , pelleted at 11,600g for 5 min, washed with 75% ethanol and vacuum dried. The DNA was dissolved in 150 μ l TE (10 mM Tris, 1 mM EDTA). DNA ($\sim 8 \mu\text{g}$ in 8 μ l) was cut with 10 U *Hae*III overnight at 37°C , then electrophoresed through a 22 cm 0.7% agarose gel for 65 h until all fragments < 3 kilobases had migrated off the gel. DNA was then dephosphorylated by treatment with 0.2 M HCl, denatured *in situ* and transferred by Southern blotting to Schleicher and Schuell BA 85 nitrocellulose membranes. High specific activity ($> 8 \times 10^8$ c.p.m. μg^{-1}) 33.6 probe was prepared by nick translation of 33.6 RF DNA with [^{32}P]dCTP and [^{32}P]dGTP (3,000 Ci mmol^{-1}). Hybridization was carried out overnight at 65°C in $1.5 \times \text{SSC}$ and $5 \times \text{Denhardt's soln.}$ The blot was washed at 65°C in $1.5 \times \text{SSC}$, 0.1% SDS. Filters were autoradiographed for four days at -80°C using two intensifying screens.

Continuing to the next generation, the resemblances between the female B and her three offspring, G, H and I are 0.415, 0.776 and 0.766. Two of these are distinctly high—lying outside the range of parent/offspring values of her previous brood. This is in accord with the incestuous nature of the mating, because a proportion of the fragments will be identical by descent. The

Individual	F	0.667								
	E	0.744	0.619							
	D	0.474	0.487	0.629						
	C	0.578	0.500	0.571	0.432					
	B	0.327	0.625	0.435	0.488	0.500				
	H	0.318	0.465	0.342	0.444	0.651	0.766			
	I	0.435	0.533	0.512	0.474	0.622	0.776	0.773		
	G	0.160	0.286	0.298	0.333	0.367	0.415	0.417	0.320	
	J	0.093	0.095	0.150	0.057	0.143	0.087	0.049	0.047	0.723
		A	F	E	D	C	B	H	I	G

Individual J is the male from the adjacent nest box that shows a high similarity to nestling G, and presumably fathered it through an extra-pair copulation.

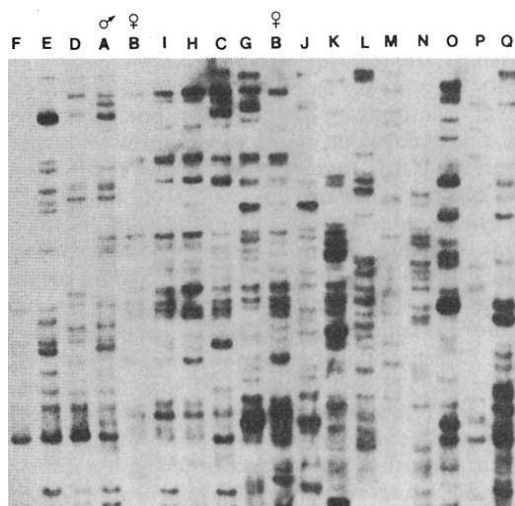


Fig. 2 Autoradiograph demonstrating the correct assignment of paternity. The pair A and B produced four offspring (C, D, E and F) in 1985. In 1986, male C mated incestuously with its mother B producing three offspring (G, H and I). G was found to be mismatched with its putative parents by starch gel electrophoresis and its fingerprint contains several bands found in neither: all these are however found in male J.

other value is much lower, suggesting that the offspring may not be incestuous. This is confirmed by comparing the offspring with their putative father (C): the similarities are 0.367, 0.622 and 0.651. Again, the latter two are high, but the first is similar in size to the value of *D* obtained from a comparison of the nestlings between broods B to E in Fig. 1, of 0.304 (s.e.m. = 0.037). This value corresponds to the similarity between half-siblings, which is the genetic relationship between C and G if they share a common mother B but are otherwise unrelated. The lack of direct parent/offspring relationship between C and G is strengthened because their enzyme bands at an isocitrate dehydrogenase locus do not match, and because they do not share any minisatellite band that could not have been inherited directly from B. More conclusively, G possesses several bands that C lacks.

We then screened birds from elsewhere in the colony in an attempt to find the true father. We chose males from the eight nearest nest boxes, and compared their banding pattern with A–H using *D*. All values were less than 0.25 with one exception: the comparison of J with G gives *D* = 0.723, well within the parent/offspring range. J was actually present in the adjacent nest box to the family in question, so that it seems probable that an extra-pair copulation took place between female B and this male. Direct comparison of G with B and J confirms the conclusion of parentage. Every band present in G occurs in one or other (or both) of its presumed parents. Assuming that the bands are inherited independently, the probability of this occurring by chance among unrelated birds is extremely low.

It is apparent from these results that the polymorphic minisatellite DNA discovered by Jeffreys *et al.* in humans and also reported in cats and dogs⁷ is equally variable in other species. It appears to be inherited in the same simple fashion and individuals are as different from one another as are people. Its use provides an excellent means for resolving relatedness in nature, and is likely to revolutionize the study of those aspects of behaviour, population genetics and biometry that require a detailed demographic knowledge of populations.

We are grateful to Dr A. J. Jeffreys for the minisatellite probe and for his helpful attitude to this research. We thank him and Professor B. C. Clarke for critically reading the manuscript, and Dr T. A. Burke for helpful discussions. We thank the Lister Institute for permission to use the human minisatellite probe, and the principal and staff of the Brackenhurst College of

Agriculture for allowing us to work at their farm. The work was funded by NERC and the Nuffield Foundation.

Received 23 January; accepted 18 March 1987.

1. Lewontin, R. C. *The Genetic Basis of Evolutionary Change* (Columbia University Press, New York 1974).
2. Jeffreys, A. J., Wilson, V. & Thein, S. L. *Nature* **314**, 67–73 (1985).
3. Burke, T. A. thesis, Univ. Nottingham (1984).
4. Jeffreys, A. J., Wilson, V. & Thein, S. L. *Nature* **316**, 76–79 (1985).
5. Bulatova, N. S., Radjabu, S. I. & Panov, E. N. *Experientia* **28**, 1369–1371 (1972).
6. Larson, R. A., Shode, R. O., Shapira, J. F. & Avise, J. C. *J. molec. Evol.* **17**, 214–226 (1981).
7. Jeffreys, A. J. & Morton, D. M. *An. Genet.* (in the press).

DNA fingerprinting in birds

T. Burke & M. W. Bruford

Department of Zoology, University of Leicester,
Leicester LE1 7RH, UK

Several regions of the human genome are highly variable in populations because the number of repeats in these regions of a short 'minisatellite' sequence varies at high frequency. Different minisatellites have a core sequence^{1,2} in common, however, and probes made up of tandem repeats of this core sequence detect many highly variable DNA fragments in several species including humans^{1,3}, cats⁴, dogs⁴ and mice⁵. The hypervariable sequences detected in this way are dispersed in the genome and their variability means that they can be used as a DNA 'fingerprint', providing a novel method for the identification of individuals^{2,6}, confirmation of biological relationships^{7,8} and human genetic analysis^{9,10}. We show here that human minisatellite-derived probes also detect highly variable regions in bird DNAs. Segregation analysis in a house sparrow family confirms that these regions comprise many mostly heterozygous dispersed loci and we conclude that house sparrow DNA fingerprints are analogous to those of humans. Fingerprint analysis identified one nestling, with fingerprint bands not present in the parent pair's fingerprints, which we conclude resulted from an extrapair copulation. Extrapair copulations have been described in many wild bird species^{11–13}, but their success and hence adaptive significance have rarely been quantifiable^{14–20}. DNA fingerprinting will be of great significance to studies of the sociobiology, demography and ecology of wild birds.

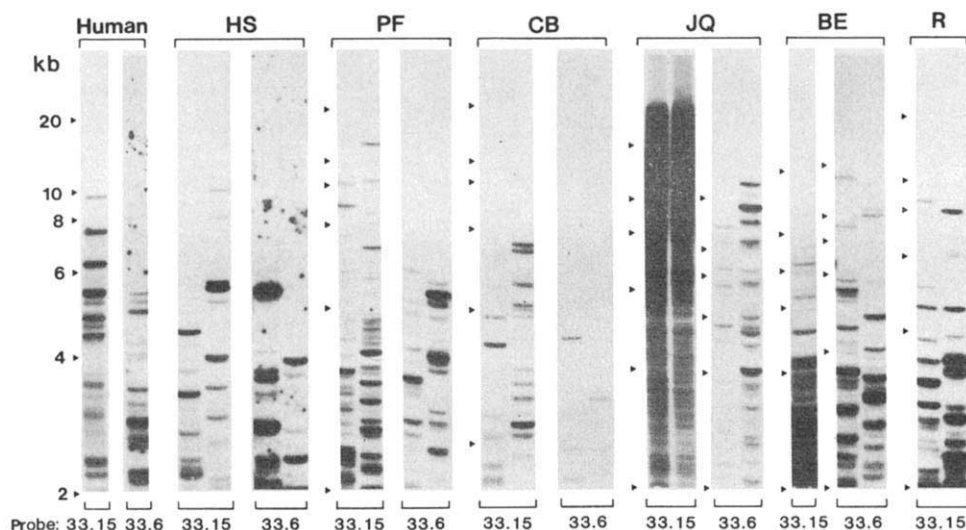
Hybridization patterns obtained from conspecifics of 16 species indicate that bird DNAs have many different and very

Table 1 Intraspecific similarities of band patterns produced by human minisatellite probes

Species	Enzyme	Probe 33.6		Probe 33.15		Combined	
		<i>n_b</i>	<i>x</i>	<i>n_b</i>	<i>x</i>	<i>n_b</i>	<i>x</i>
Human	<i>Hinf</i> I	13.8	0.21	14.7	0.21	28.5	0.21
HS	<i>Alu</i> I	6.0	0.28	15.0	0.17	27.0	0.22
PF	<i>Alu</i> I	8.0	0.13	22.0	0.27	30.0	0.23
CB	<i>Alu</i> I	5.0	0.42	15.5	0.20	20.5	0.26
JQ	<i>Hinf</i> I	17.0	0.47	—	—	—	—
BE	<i>Alu</i> I	23.0	0.30	—	—	—	—
R	<i>Alu</i> I	—	—	25.5	0.28	—	—

Probabilities of band sharing in six bird species (refer to Fig. 1 for names), with the human data (from ref. 2) for comparison. Band comparisons were made for the better resolved and less complex larger fragments only, to 4 kb for the human, PF and JQ digests, and to 2 kb for the others where the fragments are found to be of smaller average size. Here, *n_b* refers to the mean number of bands in the compared size range; *x* was found as the mean of each pairwise comparison of the proportion of individual A's bands that were matched by a band of similar intensity and electrophoretic mobility in B. A small number of weakly hybridized bands, that although apparent in A would have been obscured, if present, by a stronger band in B, were excluded. As only two individuals have been compared in each case here, confidence limits have not been estimated.

Fig. 1 Hybridization patterns obtained from avian DNAs probed with the human minisatellite probes 33.6 and 33.15, with the DNA fingerprints produced in a single human DNA for comparison. Arrowheads, estimated positions of 20, 10, 8, 6, 4 and 2 kb fragments for tracks to the right. The avian DNAs were obtained from apparently unrelated mated pairs of five species sampled in the wild (HS, house sparrow, *Passer domesticus*; PF, pied flycatcher, *Ficedula hypoleuca*; CB, corn bunting, *Miliaria calandra*; BE, European bee-eater, *Apiaster merops*; R, rook, *Corvus frugilegus*) and one laboratory population (JQ, Japanese quail, *Coturnix coturnix japonica*). DNAs were digested with restriction enzymes having a four-base recognition sequence: *HinfI* (Human, JQ) or *AluI* (HS, PF, JQ, BE, R).



Methods. Human DNA was prepared from blood². Bird tissues variously consisted of washed and lysed red blood cell fractions densely pelleted at 15,000g during an earlier isozyme survey¹⁷ and since stored for up to 6 years at -70°C (HS), frozen heparinized whole blood (BE, JQ), whole blood collected into at least one volume of $1\times\text{SSC}$ (0.15 M NaCl, 15 mM trisodium citrate, pH 7.0, autoclaved) and kept at ambient temperatures for up to 3 days before freezing at -20°C (CB, R), or pectoral muscle freshly frozen at -20°C (PF). All blood was originally collected by venipuncture. For DNA isolation, muscle tissue was first finely chopped at 0°C , cooled in liquid nitrogen and ground until particulate. About 0.5 g muscle or the equivalent of 50 μl whole blood were well suspended in 4 ml 0.1 M Tris-Cl pH 8.0, 0.1 M NaCl, 1 mM EDTA before incubation with 2.5 units ml^{-1} proteinase K (Sigma) and 0.5% SDS for at least 1.5 h at 55°C . This was followed by two phenol/chloroform and one chloroform extraction, then ethanol precipitation with 0.2 M sodium acetate. Typically, 50 μl avian blood yields $>100\text{ }\mu\text{g}$ DNA. For gels and hybridization, 5 μg samples of DNA were digested with 15 units of *HinfI* or *AluI* in the presence of 4 mM permidine trihydrochloride (Sigma) at 37°C for 16 h, and recovered after phenol extraction by ethanol precipitation, followed by washing in 80% ethanol. Digested DNAs were resuspended in a loading mix and electrophoresed in 0.7% or 1.0% (for PF only) agarose gels and blotted onto nitrocellulose filters (Schleicher and Schull, 0.45 μm pore size) as described elsewhere². The ^{32}P -labelled probe DNAs were prepared from the human minisatellite M13 recombinants 33.6 and 33.15 by primer extension^{1,2}. Southern blots were hybridized overnight at 62°C in the presence of 50 $\mu\text{g}\text{ ml}^{-1}$ denatured salmon DNA (Sigma Type III, sheared in 0.2 M NaOH, 20 mM EDTA at 100°C for 20 min, neutralized with HCl and recovered after phenol/chloroform extraction by ethanol precipitation), $2\times\text{Denhardt's}$, 0.1% SDS, 6% (w/v) polyethylene glycol 6,000 (BDH) and $1\times\text{SSC}$, and autoradiographed for 1–7 days at -70°C using Kodak X-Omat or Amersham MP film and a single Cawo intensifying screen. The autoradiographs for human and house sparrow were obtained from a single filter; hybridization to probe 33.6 was carried out after the removal, confirmed by autoradiography, of all previously hybridized 33.15 by immersing the filter in water at 85°C and allowing it to cool to 20°C .

variable regions with similarity to both of the human minisatellite poly-(core) probes 33.6 and 33.15. Figure 1 shows examples from six species (representing four oscine families and two non-passerine orders). In most cases the band patterns have a similar complexity to human DNA fingerprints (shown for comparison in Fig. 1). There was, however, one striking exception: probe 33.15 hybridized so intensively to *HinfI*-digested Japanese quail DNA that no individual bands could be readily distinguished and, unusually, hybridization was more complex for larger fragment sizes. Thus an abundant class of Japanese quail DNA, probably comprising a tandem-repetitive satellite DNA containing the occasional *HinfI* restriction site, has sequence similarity to probe 33.15.

The probability of band sharing between unrelated individuals within each species sampled from the wild is similar to that found in humans (Table 1). Greater similarity was observed in the Japanese quail, as expected from the fact that they originated from a relatively inbred laboratory population.

The genetics and germline stability of the detected DNA fragments were investigated in a house sparrow family, consisting of 2 adults and the 11 offspring raised by them in four successive broods, belonging to a closely monitored wild population¹⁷. DNA from each individual was hybridized under two different stringencies to probes 33.6 and 33.15 (Fig. 2). In nine offspring all DNA fragments were also present in either or both parents, but two nestlings were found to contain mismatched bands not present in either parent. These novel bands may have arisen by mutation, or may indicate incorrect parentage. As nestling 'X' contains many mismatched bands we conclude that X cannot be the progeny of this parental combination (Fig. 2 legend). We cannot, however, reject the mutation hypothesis

for the nestling containing a single mismatching band; many such mutant bands have been observed at mammalian minisatellite loci, and their mutation rate is known to vary significantly^{5,10}. The observed rate in house sparrows of $1/288 = 0.0035$ (see Fig. 2 legend) is close to that of $\sim 1/240$ for humans¹, and mutation is therefore a more likely explanation for a single mismatch than is incorrect parentage, even if an actual parent were a close relative of the assigned one.

The segregation of parental bands in the 10 nestlings excluding X was investigated. As with humans⁹, probes 33.6 and 33.15 detected completely different sets of bands. Some bands, identified from their electrophoretic mobility and their pattern of distribution among nestlings were detected in more than one hybridization experiment (identified by letters in Fig. 2). There were no instances of a scored (parent-specific) band occurring in all 10 offspring; thus all the bands scored were heterozygous. Although we consider it likely that some bands will be homozygous, we note that bands specific to one parent will tend to include those at the more polymorphic loci. The mean transmission frequency for 61 different parental bands was 0.52, consistent with the 1:1 mendelian segregation expected for heterozygous loci.

The cosegregation of bands in a parent was investigated by counting the number of gametic recombination events recorded in the sibship (Table 2 legend). Closely linked bands are expected to show no recombination, and allelic bands are expected to show no co-inheritance. From binomial expectation, the total number of pairwise comparisons among the 29 maternal and 32 paternal fragments that are expected to show complete linkage or allelism by chance is only 0.80 and 0.96 comparisons respectively. As many more linked and allelic combinations were

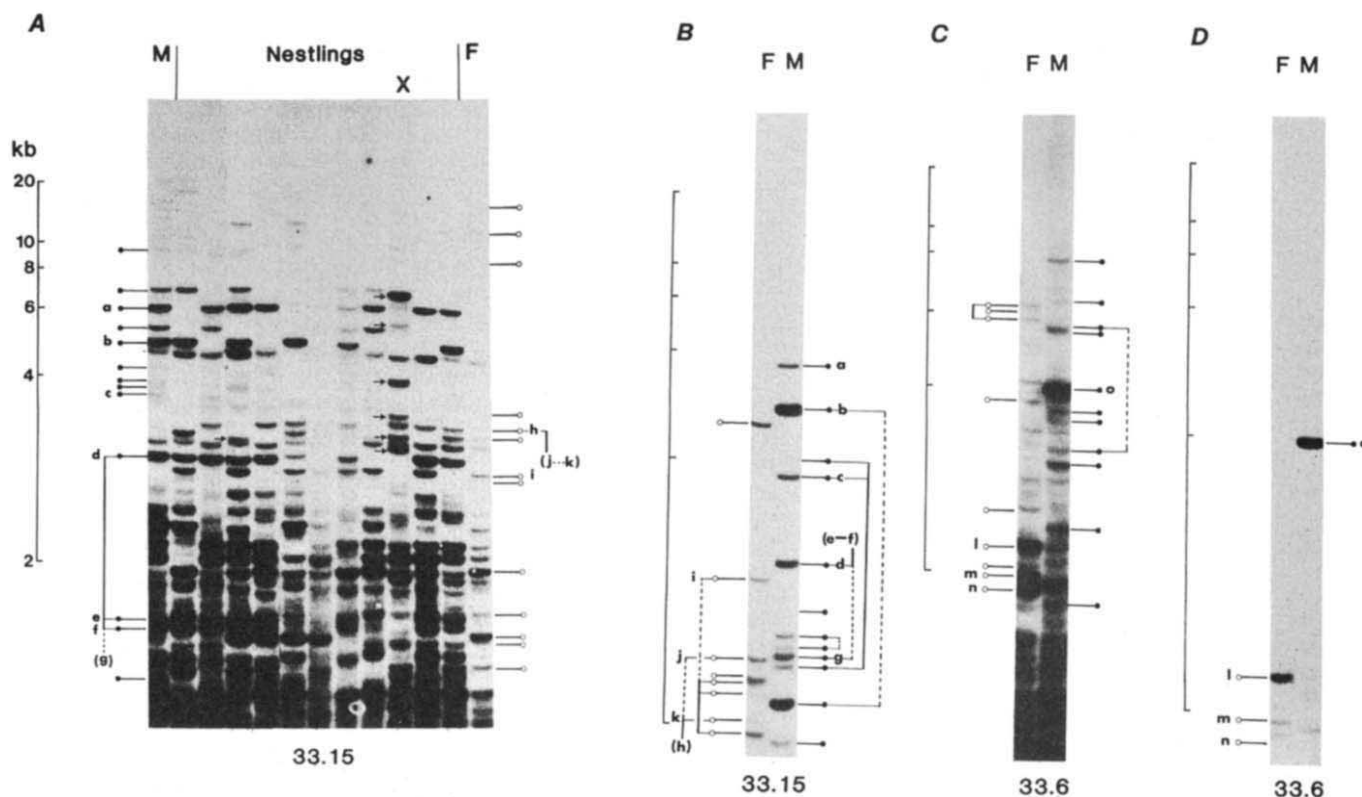


Fig. 2 Segregation of hypervariable fragments in the DNA fingerprints of a wild house sparrow family, comprising male (M) and female (F) parents with 11 nestlings raised by them in four broods, probed with: **A**, 33.15 at lower stringency; **B**, 33.15 at higher stringency; **C**, 33.6 at lower stringency and **D**, 33.6 at higher stringency. The complete sibship is shown for **A**, but only the parents for **B–D**. The broods were among those from which blood samples were obtained in a nest-box population established by us at Brackenhurst, Nottinghamshire, UK. Two nestlings contain mismatched bands (arrowed) absent from M or F. All but one of these novel bands are present in a single nestling (X). If mutation gives rise to the mismatched bands, then each mismatched band should be due to an independent mutation event, and their expected distribution among nestlings can be obtained from the Poisson distribution $e^{-m}m^n/n!$ where m is the mean number of mutant bands per individual and n is the number in a single individual. Although 7, 6, 7 and 0 mismatched bands in X were respectively observed in the experiments in **A**, **B**, **C** and **D**, we shall only consider those bands observed in **a** because bands observed elsewhere are not necessarily independent. Thus the probability of the observed 6 of a total 7 novel bands occurring in 1 of 11 offspring is $0.0000488 \times 11 = 0.000537$, enabling rejection of the mutation explanation for the mismatched bands in X. Resolved paternal (●) and maternal (○) heterozygotes are indicated. Linked pairs PQ which segregate PQ or — into offspring (excluding X) are joined by continuous lines; allelic pairs PQ segregating P- or -Q are joined by broken lines. Some fragments could be identified in both **A** and **B** or **C** and **D**, and are identified by letters.

Methods. DNA was cut with *AluI*, electrophoresed in a 30 cm 1.0% agarose gel, Southern blotted and hybridized with probes 33.6 and 33.15 (see Fig. 1 for details). Samples from both parents were run at each side of the sibship to facilitate scoring (not shown). Lower stringency hybridizations (**A**, **C**) were carried out in $1.3 \times \text{SSC}$ at 62°C and higher stringency hybridizations (**B**, **D**) in $1 \times \text{SSC}$ at 64°C , otherwise as for Fig. 1. The segregation of fragments among offspring was scored directly from the original autoradiographs taken at short and long exposures. DNA fragments were scored only when their position and intensity in each offspring matched those in the parent. For mutation rate estimation, the fragment sample size was based on the total number of clearly resolved bands (to 2 kb) found in the nestlings excluding X. For bands to 2 kb at the lower stringency conditions used here, for 33.6: $x = 0.18$, $n_b = 16$; for 33.15: $x = 0.24$, $n_b = 24$; combined $x = 0.21$. The probability of individuals other than identical twins having identical fingerprints (see text) was found as described (Table 1 legend in ref. 4). Parentage was analysed as follows. For haplotype analysis, the number of parent-specific haplotypes present in X was compared with expectation; for the 25 haplotypes of M, the expected value is 12.5, and the observed is 2 ($G_i = 20.7$, $P < 0.001$, therefore M is not the father of X); for the 19 haplotypes of F, the expected is 9.5 and the observed is 12 ($P > 0.1$). In band sharing analysis, if F were a non-relative it would be expected to share x of X's 38 resolved bands (observed 22, expected 8.1, $P < 0.001$, therefore X and F must be related). If F were X's mother, they would be expected to share 0.58 of each other's bands from which the calculated expected of 22.1 is close to the observed 22 (parent-offspring band sharing is coincidentally equal to sib band sharing when identical bands are non-allelic as calculated in ref. 7, J. F. Y. Brookfield, personal communication). The expected band sharing between X and its mother's sister will be $1 - (1 - ((1+x)/(2(2-q))))(1-q) = 0.396$ (where q is the estimated mean allele frequency, ref. 21 and J. F. Y. Brookfield, personal communication). That F is actually X's aunt can therefore be excluded ($P > 0.05$).

actually found (Fig. 2), we conclude that most cosegregating bands represent DNA fragments in close linkage, possibly single minisatellites that have been cleaved into two or more fragments, and that most apparently allelic bands represent true alleles.

The number of recombination events occurring in the sibship between each pair of loci was compared with that expected assuming independent assortment (Table 2). As the observed distributions are generated from pairwise and hence non-independent comparisons the usual goodness-of-fit tests are inappropriate, but inspection reveals no suggestion of clumping. We conclude that the house sparrow loci scored here are disper-

sed throughout the genome, as are similar loci in the four mammalian species that have so far been investigated^{4,5,9}.

The mean number of loci scored in a human DNA fingerprint (29.5, using both probes)⁹ is two to three times that for cats (10.0) and dogs (13.0) (ref. 4). The mean number scored in house sparrows at the lower stringency (21.5) is intermediate, whereas few (7.5) are detected at the higher stringency (Fig. 2). Alleles in the higher stringency fingerprints are represented by an average 1.20 fragments, whereas the corresponding value for the lower stringency is 1.07 and that for humans is only 1.01. Thus there is greater independence among bands in the more

Table 2 Segregation analysis for the house sparrow family

Transmission to no. of offspring (r)	Female				Male			
	Single fragment		Transmission of Pair (PQ or --)		Single fragment		Transmission of Pair (PQ or --)	
	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.
0	0	0.0	0	0.2	0	0.0	0	0.3
1	0	0.2	2	2.5	1	0.2	5	2.7
2	1	1.0	13	11.1	0	1.1	11	12.1
3	4	2.7	30	29.6	2	2.8	24	32.3
4	5	4.7	55	51.9	6	4.9	56	56.6
5	4	5.7	60	62.3	5	5.9	66	67.9
6	3	4.7	41	51.9	5	4.9	55	56.6
7	4	2.7	36	29.6	2	2.8	44	32.3
8	1	1.0	14	11.1	2	1.1	11	12.1
9	1	0.2	2	2.5	1	0.2	4	2.7
10	0	0.0	0	0.2	0	0.0	0	0.3
Total	23		253		24		276	
Transmission frequency $\pm$ s.e.		0.521 $\pm$ 0.037				0.509 $\pm$ 0.038		

Segregation of hypervariable loci in the house sparrow family (excluding offspring X). All but one of each set of linked or pair of allelic fragments in a parent (Fig. 2) have been excluded, so that each scored locus is included once only. The 23 maternal and 24 paternal loci are represented by 29 and 32 fragments respectively. Linkage was scored according to the number of offspring concordant for the presence (PQ) or absence (--) of each pairwise combination of loci identified in a parent. By definition, no pairs fall into the 0- or 10-offspring classes representing alleles and linked pairs respectively. Obs, observed; exp, expected values: expected binomial distributions (see ref. 9) assume transmission frequencies of 50% and independent assortment. Sex-linkage could not be investigated as the sexes of nestlings were unknown.

complex of the sparrow fingerprints, possibly because many linked fragments occur in the unresolved region, but house sparrow bands are less independent than those of humans. This may suggest that the sparrow fragments are more often clustered into tight linkage groups than in humans, or that *AluI* (as used for sparrow fingerprints) cleaves fragments internally more frequently than *HinfI* (as used for humans).

Having demonstrated the independence of most bands in the lower stringency fingerprints, we can estimate the theoretical probability that two individuals will share the same pattern from a knowledge of the frequency with which bands of similar mobility are shared. For probe 33.15, this probability is $\sim 2 \times 10^{-20}$ for two random unrelated individuals and $\sim 9 \times 10^{-10}$ for two sibs; for probe 33.6 the respective probabilities are $\sim 3 \times 10^{-14}$ and $\sim 5 \times 10^{-7}$ (Fig. 2 legend).

Knowledge of the independence of bands is also required in the analysis of relatedness. We wish to know if either adult is a parent of nestling X (see Fig. 2). As we have identified linked and allelic bands we can consider the alleles, or 'haplotypes' comprising closely linked bands, at the independent loci detected at either or both stringencies. As all the scored haplotypes are heterozygous, we expect half to occur in an offspring. As X has fewer than expected of the haplotypes of the male member of the pair (M) ($P < 0.001$), we conclude that M was not X's biological parent. The female (F) cannot be similarly excluded ($P > 0.1$). We also note that where both alleles have been detected at a locus, one of each allelic pair should occur in each offspring: whereas X has one of each of F's two identified pairs, it has only one of four expected from M. Further comparisons have to be based on band-sharing rather than the inheritance of identified haplotypes, and will be restricted to the fragments resolved in the lower stringency hybridizations, as these appear reasonably independent. Comparison of band sharing between X and F with that expected if they were unrelated clearly demonstrates that they must be related ($P < 0.001$), and we can exclude the possibility that X's true mother was a close relative

of F, such as a female sibling of F ($P < 0.05$; see Fig. 2 legend). We cannot exclude the possibility that X is actually F's full sibling, as could occur if F's parents deposited an egg in F's nest, but we consider this unlikely as there was no direct evidence for intraspecific nest parasitism in this study population¹⁷ or for the proximity of close relatives. We conclude that X is most probably the result of an extrapair copulation between F and an unidentified male.

Using a single probe (33.15 at the lower stringency) we estimate that the probability of non-detection of an incorrectly assigned father will be $\sim 3 \times 10^{-6}$ if it is unrelated to the actual father and ~ 0.01 if it is a close relative (brother or father)². This corresponds to an exclusion probability of > 0.99 , which compares with a probability of exclusion for seven blood enzyme loci sampled in the same house sparrow population of only 0.51 (ref. 17).

The term 'DNA fingerprint' was chosen because of the predicted individual specificity of human minisatellite band combinations. Although we have shown that analogous DNA fingerprints can be obtained in the house sparrow, we consider that the patterns obtained in other species should not be regarded as true DNA 'fingerprints' until genetic analysis has demonstrated that the component bands represent dispersed loci. Also, this should be done for each set of reaction conditions: the degree of independence among bands is likely to vary as their number changes. For example, we found that many bands belonged to only a few loci when their number was limited in hybridizations of higher stringency (Fig. 2). Nonindependence might also increase as more DNA fragments hybridize under less stringent reaction conditions. The result of probing Japanese quail DNA with probe 33.15 strongly suggests that sequences other than minisatellites may be detected by these probes.

In addition to its many other uses, DNA fingerprinting will clearly be of crucial value for the study of sexual selection, mating behaviour and demography of wild birds and mammals, but the full power of the method will be most readily realized in species where the genetic analysis of large families is possible.

We thank Alec Jeffreys and his colleagues Ila Patel, Vicky Wilson and Zilla Wong for their advice and interest. Samples were provided by M. I. Avery and J. R. Krebs (bee-eaters), P. Bateson (quail), E. S. Lawrence (rooks), and A. Lundberg (flycatchers). We also thank J. F. Y. Brookfield, A. Tharme, D. B. A. Thompson, J. Cutteridge, G. P. Wallis, L. Baldwin, H. C. Macgregor, L. Barnett, E. Wallace, J. Brewin and S. Burke. The corn bunting samples were obtained with the support of an Association for the Study of Animal Behaviour small grant to T. B. and D. B. A. Thompson. T.B. is a research fellow of the NERC. This work was supported by an SERC grant to T. B. and H. C. Macgregor.

Received 24 February; accepted 13 March 1987.

- Jeffreys, A. J., Wilson, V. & Thein, S. L. *Nature* **314**, 67-73 (1985).
- Jeffreys, A. J., Wilson, V. & Thein, S. L. *Nature* **316**, 76-79 (1985).
- Wong, Z. & Jeffreys, A. J. *Nucleic Acids Res.* **14**, 4605-4616 (1986).
- Jeffreys, A. J. & Morton, D. B. *Anim. Genet.* **18** (in the press).
- Jeffreys, A. J., Wilson, V., Kelly, R., Taylor, B. A. & Bullfield, G. *Nucleic Acids Res.* (in the press).
- Gill, P., Jeffreys, A. J. & Werrett, D. J. *Nature* **318**, 577-579 (1985).
- Jeffreys, A. J., Brookfield, J. F. Y. & Semeonoff, R. *Nature* **317**, 818-819 (1985).
- Hill, A. V. S. & Jeffreys, A. J. *Lancet* **i**, 1394-1395 (1985).
- Jeffreys, A. J., Wilson, V., Thein, S. L., Weatherall, D. J. & Ponder, B. A. J. *Am. J. hum. Genet.* **39**, 11-24 (1986).
- Jeffreys, A. J. *Biochem. Soc. Trans.* (in the press).
- Gladstone, D. E. *Am. Nat.* **114**, 545-557 (1979).
- Ford, N. L. *Current Ornithology* **1**, 330-356 (1983).
- McKinney, F., Cheng, K. M. & Bruggers, D. J. in *Sperm Competition and the Evolution of Animal Mating Systems* (ed. Smith, R. L.) (Academic, New York, 1984).
- Burns, J. T., Cheng, K. M. & McKinney, F. *Auk* **97**, 875-879 (1980).
- Cheng, K. M., Burns, J. T. & McKinney, F. *Auk* **100**, 302-310 (1983).
- Alatalo, R. V., Gustafsson, L. & Lundberg, A. *Oikos* **42**, 41-47 (1984).
- Burke, T. thesis, Univ. Nottingham (1984).
- Gavin, T. A. & Bollinger, E. K. *Auk* **102**, 550-555 (1985).
- Westneat, D. F. *Anim. Behav.* (in the press).
- Quinn, T. W., Quinn, J. S., Cooke, F. & White, B. N. *Nature* **326**, 392-394 (1987).
- Hill, W. G. *Nature* **322**, 290-291 (1986).

Inner ear of the coelacanth fish *Latimeria* has tetrapod affinities

B. Fritzsch

University of Bielefeld, Faculty of Biology, 4800 Bielefeld 1, FRG

Auditory reception in elasmobranchs, teleosts and amphibians may be mediated by various inner-ear sensory epithelia¹⁻³, including the basilar papilla, which seems to be the precursor of the cochlea in mammals. The origin of the basilar papilla remains a major unsolved problem for understanding the evolution of hearing in terrestrial vertebrates⁴⁻⁶. Study of living species indicates that the basilar papilla is a unique feature of tetrapods^{6,7}, but palaeontological data indicate that this epithelium as well as a middle ear, is already present in crossopterygian fish⁸⁻¹⁰. However, no basilar papilla has been found in the only living crossopterygian species, the coelacanth *Latimeria chalumnae*¹¹. I have re-examined the inner ear of adult and embryonic *Latimeria* and find a membranous specialization which resembles in structure, position and innervation pattern the basilar papilla of tetrapods, in particular amniotes. No epithelium comparable to the basilar papilla was found in lungfish. I suggest that the basilar papillae of *Latimeria* and tetrapods are homologous and evolved only once in their common ancestor.

For this study the inner ear of an adult male *Latimeria* (CAS, 33111) was dissected out and stained for myelinated nerve fibres with 0.5% Sudan black B in 70% ethanol. Sections of an embryonic *Latimeria* (AMNH, 197650) embedded in nitrocellulose, cut at 15 μ m thickness and stained with trichrome stain, were also available. Because of the recently disputed phylogenetic affinity of *Latimeria* and lungfish to tetrapods¹², the inner ear of *Protopterus aethiopicus* (CAS 51379 and 66381) and *Lepidosiren paradoxa* (CAS 36480) were also re-examined. No differences from earlier descriptions⁴ were detected, and therefore previous descriptions^{4,13,14} of the inner ear of *Neoceratodus*

forsteri were used without re-examination of specimens. I also examined the inner ear of a reptile (*Gekko gekko*) for comparison.

The cartilage of the otic capsule, the configuration of the labyrinth, and the major nerve tracts largely confirm a previous description of *Latimeria*¹¹ but differ in three major aspects. First, the utricular macula lies at the base of the semicircular canals, as in tetrapods and actinopterygian fishes and not within an utricular recess as in Chondrichthyes and Dipnoi^{4,5} (Fig. 1). Most fibres to the utricular macula run in a separate medial twig, but some run together with the fibres to the anterior and horizontal canal, as in some mammals⁵. A separate lagenar recess including the lagenar epithelium was found in *Latimeria* (Fig. 1a), in contrast with lungfish where the lagenar epithelium lies within the saccular chamber (Fig. 1b) as it does in chimaeras, chondrosteans, holosteans and some teleosts^{2,4,13}.

Second, the posterior eighth nerve branch sends one ramus to the posterior canal, one to the lagena, and several rami to the sacculus. The posterior canal ramus gives rise to three twigs which run to a single epithelial patch covered with a cupula in the posterior canal near the utriculo-saccular foramen (Figs 1 and 2). One or two sensory epithelia covered with a cupula are known in all vertebrate classes in a comparable position and with a comparable fibre supply^{1,2,4,6}. These similarities in position, nerve supply and the presence of a covering cupula suggest that this epithelium is homologous to the papilla neglecta of other vertebrates (Fig. 2). We found no evidence of an amphibian papilla, believed to represent a part of the neglected papilla in amphibians^{4,6}.

Third, the ramus to the posterior canal and the papilla neglecta has a few fibres forming two bundles to the lagena and to a foramen in the wall of the sacculo-lagenar orifice. Only anurans have a nerve supply to their basilar papilla which follows a similar course⁶. Most lagenar fibres travel in a ramus which sweeps ventro-medially around the lagenar recess. Near the foramen at the sacculo-lagenar orifice many fibres leave the lagenar twig and run in a widespread, thin fibre layer to the foramen (Figs 1, 2). All fibres lose their myelin sheaths at this

Fig. 1 The organization of the left labyrinth of the coelacanth *Latimeria* (a) and of the lungfish *Protopterus* (b) viewed from a medial aspect. Several nerve twigs run to the neglected papilla (NP). In *Latimeria* the utriculus (U) receives fibres from a medial twig and the sacculus (S) from the posterior ramus only. The lagena (L) receives fibres through a lagenar branch and the posterior vertical canal branch. The lagenar ramus also carries fibres to the basilar papilla (BP) at the sacculo-lagenar orifice. In lungfish (b) the utricular macula (U) is within a separate utricular recess. No lagenar recess is present and the lagenar epithelium (L) shares a recess with the saccular macula (S). No basilar papilla is present in lungfish.

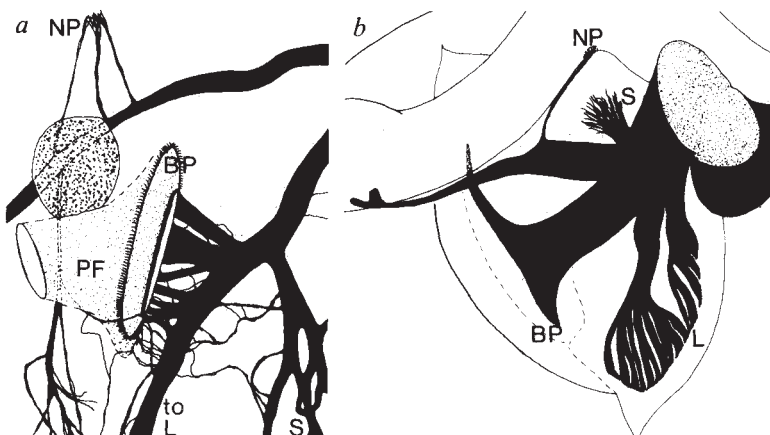
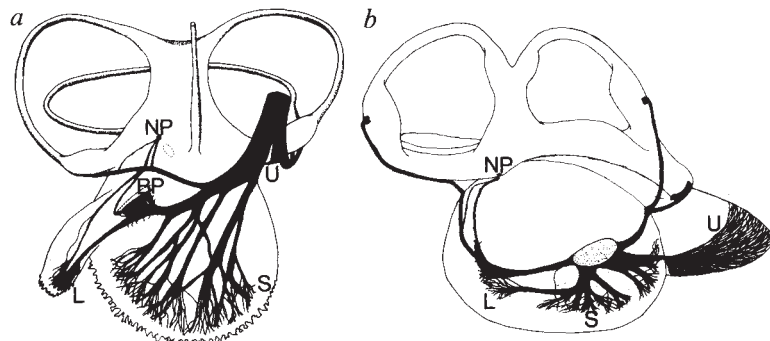


Fig. 2 The fibre supply to the basilar papilla (BP) of *Latimeria* (right ear, a) and *Gekko gekko* (b). In both animals these fibres separate from the lagenar fibres (L). Fibres to the neglected papilla (NP) separate from the ramus to the posterior canal. The perilymphatic foramen (PF) is indicated by fine stippling, cartilage of the ear capsule by coarse stippling. Magnification, $\times 4$.

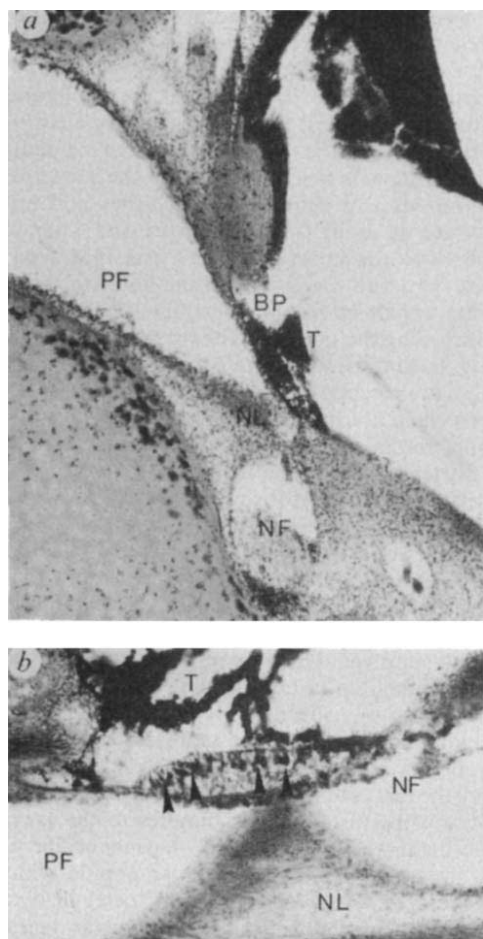


Fig. 3 *a*, The basilar papilla (BP) separates the sacculus (S) from the perilymphatic foramen (PF) in the otic capsule of *Latimeria*. *b*, Nerve fibres (NF) course to the basilar papilla and the tissue of the neural limbus (NL) partly supports the basilar papilla as in amniotes². A tectorial structure (T) overlies the hair cells (arrowheads). Magnification, $\times 140$ in *a*; $\times 190$ in *b*.

foramen. Both in position and in trajectory, these nerve fibres resemble the nerve supply to the basilar papilla of amniotes (Fig. 2)^{2,4}.

A membrane covers the foramen at the sacculo-lagenar orifice. In the inner ears dissected out of *Latimeria*, this membrane is present only as a thin flap of tissue at the side of the foramen. In the sectioned material of the embryo, the foramen is sealed by a thin membrane which is studded with sensory cells at the saccular luminal side (Fig. 3) and is partly supported by the neural limbus as in amniotes². This membrane suggests that the canal connecting both inner ears is a perilymphatic as in tetrapods⁵ rather than an endolymphatic duct as in some teleosts^{1,2} and as previously suggested¹¹. Dark-staining fibrils above the basilar papilla may represent a tectorial structure. The dimensions ($\sim 5.8 \text{ mm} \times 0.9 \text{ mm}$) and structure of the membrane are comparable to those of the basilar papilla of amniotes and the utricular macula of clupeids^{1,2}. As in the latter cases, the membrane of *Latimeria* might be responsive to very small pressure changes between the endolymphatic cavity and the perilymphatic duct. The way such pressure changes might be achieved is at present unclear: no accessory structure important for aquatic hearing¹ has been reported for *Latimeria*¹¹.

In addition to the clupeid utricular macula there is only one other sensory epithelium in the vertebrate inner ear that structurally resembles the epithelium of *Latimeria* described here: the tetrapod basilar papilla. Because it is different in both position and fibre supply^{1,2}, the clupeid utricle is clearly only a functional analogue of both the basilar papilla of tetrapods

and the sensory epithelium of *Latimeria*. The latter two organs, however, have a number of similarities: both are located at the lagenar-sacculus orifice, that is, they are homotopic; in both the fibre supply is from a lagenar twig (as in amniotes, urodeles, gymnophionans and *Latimeria* Figs 1 and 2) or from the posterior canal twig (as in anurans and *Latimeria*); and in both, other than in amphibians, the basilar hair cells rest largely on a membrane suspended between an endo- and a perilymphatic space (Fig. 3), a structural similarity which implies a similar function.

These considerations suggest that the sensory epithelium of *Latimeria* should be classified as a basilar papilla. Whether it is homologous to the tetrapod basilar papilla or represents a case of parallel evolution is not clear. However, if we assume that the basilar papillae of *Latimeria* and tetrapods formed twice, in separate lineages, we must also assume that the basilar papilla of amphibians has evolved in parallel to the basilar papillae of both *Latimeria* and amniotes. This conclusion is necessary because of the larger structural differences between the amphibian basilar papilla and both the *Latimeria* and amniotic basilar papilla, and implies that the basilar papilla has evolved three times independently.

There is another argument in favour of homology of the basilar papilla of *Latimeria* and of tetrapods. In both fossil coelacanths and rhipidistians there is a duct between the two inner ears^{8,15}. As in *Latimeria* and tetrapods, including humans⁵, this duct might be a perilymphatic duct originating on the perilymphatic foramen adjacent to the basilar papilla, but an endolymphatic nature for this duct, as in some teleosts^{1,2}, cannot at present be ruled out. Note that no structure resembling the basilar papilla is present in lepidosirenid lungfishes (Fig. 1) where the lagena is not even fully separated from the sacculus^{4,13,14}. Because rhipidistians are generally considered ancestors of tetrapods, this would indicate that evolution of the basilar papilla began more than 350 million years ago in water rather than on land as previously suggested^{6,10}. Moreover, the hyomandibula of *Latimeria* (which becomes the stapes in the tetrapod middle ear⁷) serves as a jaw suspension element¹¹ rather than being connected to a round window as in tetrapods. Likewise, the brainstem shows as in urodeles, but opposite to anurans or amniotes¹⁶, the non-teleostean pattern^{3,17} without recognizable auditory nuclei. Thus it appears likely that the early differentiation of the basilar papilla was linked with the differentiation of the perilymphatic system but not with either middle-ear formation or central nervous system specialization.

I thank the staff of the California Academy of Sciences for permission to dissect a *Latimeria*, Dr M. Lagios and Mrs S. Concannon for kindly providing the sections of an embryo, the Deutsche Forschungsgemeinschaft for financial support and Drs M. H. Wake, W. Wilczynski, J. A. Winer and A. Bauer for comments on earlier drafts of the manuscript.

Received 23 January; accepted 12 March 1987.

1. Tavolga, W. N., Popper, A. N. & Fay, R. R. *Hearing and Sound Communication in Fishes* (Springer, New York, 1981).
2. Lewis, E. R., Leverenz, E. L. & Bialek, W. S. *The Vertebrate Inner Ear* (CRC Press, Boca Raton, Florida, 1985).
3. Northcutt, R. G. in *Comparative Studies of Hearing in Vertebrates* (eds Popper, A. N. & Fay, R. R. 75-118 (Springer, New York, 1980).
4. Retzius, G. *Das Gehörorgan der Wirbeltiere* (Samson & Walin, Stockholm, 1881).
5. Werner, C. F. *Das Gehörorgan der Wirbeltiere und des Menschen* (Thieme, Leipzig, 1960).
6. Baird, I. L. in *Handbook of Sensory Physiology*, Vol. V/1, (eds Keidel, W. D. & Neff, W. D.) 159-212 (Springer, Berlin, 1974).
7. Lombard, R. E. & Bolt, J. *Biol. J. Linn. Soc.* **11**, 19-76 (1979).
8. Romer, A. S. *Bull. Mus. comp. Zool. Harv.* **82**, 1 (1937).
9. Jarvik, E. *Basic Structure and Evolution of Vertebrates* (Academic, London, 1980).
10. Van Bergeijk, W. A. *Am. Zool.* **6**, 371-377 (1966).
11. Millot, J. & Anthony, J. *Anatomie de Latimeria chalumnae* Vol. 2 (Centre National Recherche Scientifique, Paris, 1965).
12. Rosen, D. E., Forey, B. G., Gardiner, B. G. & Patterson, C. *Bull. Am. Mus. nat. Hist.* **167**, 6163-275 (1981).
13. Burns, R. H. *Anat. Anz.* **43**, 396-400 (1913).
14. Gaudie, R. W., Dunlop, D. & Tse, J. *N.Z. J. mar. Fresh. Res.* **20**, 81-92 (1986).
15. Bjerring, H. C. in *Evolutionary Biology of Primitive Fishes* (eds Gorbman, A., Dodd, J. M. & Olsson, R.) 31-57 (Plenum, New York, 1985).
16. Fritsch, B., Nikundiwe, A. M. & Will, U. *J. comp. Neurol.* **229**, 451-468 (1984).
17. Kremers, J. W. P. M. & Nieuwenhuys, R. *J. comp. Neurol.* **187**, 613-638 (1979).

Substance P-immunoreactive retinal ganglion cells and their central axon terminals in the rabbit

N. Brecha*†‡§||, D. Johnson†, J. Bolz¶,
S. Sharma#, J. G. Parnavelas**
& A. R. Lieberman**

* Department of Medicine, † Department of Anatomy,
‡ Jules Stein Eye Institute, § Brain Research Institute,
|| Center for Ulcer Research and Education, UCLA School of
Medicine and VA Medical Center-Wadsworth, Los Angeles,
California 90073, USA

¶ Max Planck Institut für Entwicklungsbiologie,
D-7400 Tübingen 1, FRG

Department of Ophthalmology, New York Medical College,
Valhalla, New York 10595, USA

** Department of Anatomy and Embryology, University College London,
Gower Street, London WC1E 6BT, UK

Retinal ganglion cells are the projection neurons that link the retina to the brain. Peptide immunoreactive cells in the ganglion cell layer (GCL) of the mammalian retina have been noted but their identity has not been determined¹⁻³. We now report that, in the rabbit, 25–35% of all retinal ganglion cells contain substance P-like (SP) immunoreactivity. They were identified by either retrograde transport of fluorescent tracers injected into the superior colliculus, or by retrograde degeneration after optic nerve section. SP immunoreactive cells are present in all parts of the retina and have medium to large cell bodies with dendrites that ramify extensively in the proximal inner plexiform layer. Their axons terminate in the dorsal lateral geniculate nucleus, superior colliculus and accessory optic nuclei, and these terminals disappear completely after contralateral optic nerve section and/or eye enucleation. In the dorsal lateral geniculate nucleus large, beaded, immunoreactive axons and varicosities make up a narrow plexus just below the optic tract, where they define a new geniculate lamina. The varicosities make multiple synaptic contacts with dendrites of dorsal lateral geniculate nucleus projection neurons and presumptive interneurons in complex glomerular neuropil. This is direct evidence that some mammalian retinal ganglion cells contain substance P-like peptides and strongly suggests that, in the rabbit, substance P (or related tachykinins) may be a transmitter or modulator in a specific population or populations of retinal ganglion cells.

New Zealand and Dutch Belted rabbits (not operated upon or surviving 3–143 days after unilateral optic nerve section or eye enucleation) were used. Most received an intraocular injection of colchicine, which is known to increase peptide levels in neuronal cell bodies⁴. Some received a unilateral injection of Fast Blue⁵ or rhodamine-labelled microspheres⁶ into the superficial layers of the superior colliculus (SC) to label ganglion cells, all or most of which are known to project to the SC in this species⁷. Wholemounds of retina or sections of retina or brain were processed by standard immunohistochemical techniques^{8,9}, using a monoclonal antibody against substance P (see ref. 10, and Fig. 2 legend for technical details and specificity controls). Although control experiments cannot completely exclude the possibility of cross-reactivity with other substances, including related tachykinins¹¹, the terms 'SP-immunoreactive' or 'SP-containing' are used to describe stained tissue.

SP immunoreactive cell bodies and processes were present in all regions of both colchicine-treated and untreated retinas but not in retinas incubated with antibodies absorbed with substance P. A minority of immunoreactive cells were in the inner nuclear layer (Fig. 2*a, b*). In keeping with previous reports^{1,2,12} some were unistratified amacrine cells which ramified in lamina 5 (ref. 13) whereas other slightly larger cells ramified in laminae 1, 3 and 5 of the inner plexiform layer and possibly also in the outer

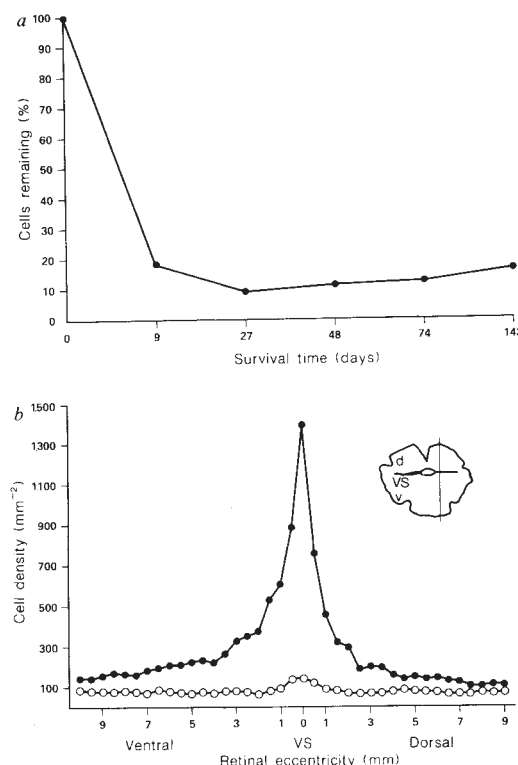


Fig. 1 *a*, Number of SP-immunoreactive cells remaining in the GCL of optic-nerve-sectioned compared to normal (contralateral) retina. Counts were made over the visual streak near the centre of the retina. *b*, Distribution and density of SP immunoreactive cells in optic-nerve-sectioned (open circles) compared to normal (contralateral) retina (solid circles; 74 day survival period). VS, visual streak.

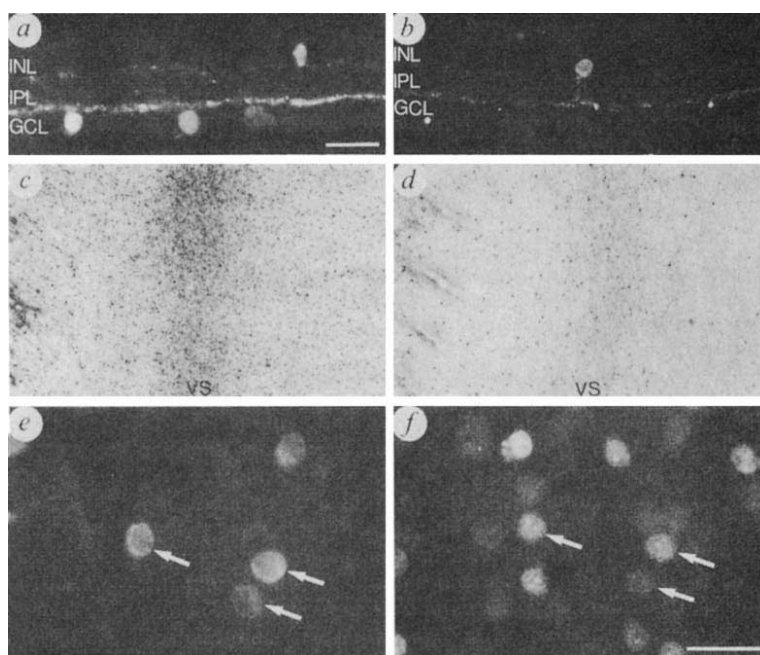
Methods. The number of SP-immunoreactive somata was determined from wholemounts in two adjacent 0.5-mm² fields at 0.5 mm intervals (final magnification $\times 272$) along a dorso-ventral axis perpendicular to the visual streak and located about 3 mm from the optic nerve head (see inset; d, dorsal; v, ventral, VS visual streak). The average number of cells of the adjacent fields and cell density was calculated for each interval from these maps. Cell density is probably underestimated in the region of the myelinated fibre and retinal blood vessel band (retinal eccentricity about 2–4 mm dorsal to the visual streak) owing to difficulties in visualizing stained somata. Measurements were not corrected for retinal shrinkage.

plexiform layer (multistratified amacrine and/or interplexiform cells).

Most of the immunoreactive perikarya were, however, situated in the GCL (Fig. 2*a, b*), and typically gave rise to three or four dendrites which branched extensively in the proximal inner plexiform layer (lamina 5), adjacent to the GCL. Primary and secondary dendrites could be followed for up to 100 μ m from their parent cell bodies before they were lost among other stained processes. The absolute dimensions of the dendritic fields of individual cells could not therefore be determined. Dendrites of neighbouring cells overlapped extensively so that any given region of the retina was in the fields of several cells. Many of these cells also gave rise to an axon which could be followed through the nerve fibre layer towards the optic nerve head. Although there were some medium-large multipolar cell bodies, most immunoreactive cells in the GCL were ovoid in shape and had diameters between 9 and 12 μ m (for central retina the mean diameter was 10.2 ± 1.3 ($N = 257$); for the peripheral retina, 11.6 ± 2.5 ($N = 113$)) which is within the reported range of diameters of ganglion cells and the largest displaced amacrine cells^{7,14-18}. The immunoreactive cells of the GCL occur in all parts of the retina with a distribution density that parallels the

Fig. 2 Localization of SP-immunoreactive somata and processes in the rabbit retina. Immunoreactive somata are present in the GCL as well as the inner nuclear layer (INL). Processes are distributed to laminae 1, 3 and 5 of the inner plexiform layer (IPL) with the highest density in lamina 5 where they form a prominent continuous plexus. *a, b*, Transverse cryostat sections of the retina through the visual streak (retinal region with highest cell density) showing SP-immunoreactive cells in the GCL and INL. *c, d*, Wholemount preparations showing SP-immunoreactive somata in the GCL of the visual streak (VS) in a normal retina (*c*) and marked loss of immunoreactive cells 74 days after optic nerve section (*d*). *e, f*, Fast Blue retrogradely labelled SP immunoreactive ganglion cells near the visual streak in a wholemount. *e*, SP immunoreactive somata in the GCL. *f*, Same field as in *e* illustrating Fast Blue fluorescent somata. Arrows in *e* and *f* show the same cells containing both SP-immunoreactivity and Fast Blue fluorescence. Scale bars, 25 μ m. Figs. 1c, d, $\times 20$.

Methods. For intraocular colchicine treatment, rabbits ($N = 15$) were anaesthetized and 10–200 μ g of colchicine (Sigma) in 50 μ l of 0.9% NaCl was injected into the eye, just posterior to the ora serrata, 12–48 h before perfusion. Rabbits were subsequently anaesthetized and perfused transcardially with 0.9% NaCl in 0.1 M phosphate buffer (pH 7.4). For retinal sections, the NaCl perfusion was followed by 4% paraformaldehyde, the eyes were removed, the anterior segments cut away and the eyecups postfixed. Eyecups were subsequently transferred and washed in 20% sucrose, serial sections were cut perpendicular to the vitreal surface and processed for immunohistochemistry. For retinal wholemount preparations the retina was dissected from the pigment epithelium, fixed, washed in 20% sucrose and subsequently frozen, thawed and processed for immunohistochemistry. Wholemounts or sections were incubated for 12–96 h at 4 °C with a rat monoclonal antibody¹⁰. For immunofluorescence procedures, the tissue was subsequently washed, incubated in anti-rat immunoglobulin G (IgG) conjugated to either fluorescein isothiocyanate (FITC) or tetramethylrhodamine (TRITC) (Accurate), washed, mounted (GCL side up for retinal wholemounts) and covered with a cover slip. For peroxidase-antiperoxidase and avidin-biotin procedures, the tissue was washed, incubated in anti-rat IgG or biotinylated IgG (Accurate or Vector), washed, incubated in peroxidase-antiperoxidase or avidin-biotin-peroxidase complex (Accurate or Vector), washed and then incubated in 3',3'-diaminobenzidine solution with 0.05% H_2O_2 and washed. Tissue was mounted (GCL side up for wholemounts), incubated in 0.01% OsO_4 , washed, cleared and covered with a cover slip. Specificity was assessed by substituting the primary antibody with normal rat serum (1:10) or primary antibody which had been previously absorbed overnight at 4 °C with 10 μ M synthetic substance P (Balchem). For optic nerve section, rabbits ($N = 20$) were deeply anaesthetized with a mixture of xylazine and ketamine and a 1-cm incision was made away from the lateral canthus, the eyelid gently retracted and the eye pulled slightly forward to expose the optic nerve. The optic nerve was sectioned, the skin sutured and the animal was allowed to recover and survive for 2–143 days. The retinas and brain were then fixed and processed as above for immunohistochemistry. For Fast Blue and rhodamine-labelled microsphere injections, rabbits ($N = 8$) were deeply anaesthetized, the skull and cortex overlying the SC was removed and an injection of 5 μ l of 3% Fast Blue or 5 μ l of rhodamine labelled microspheres (diameter 0.02–0.2 μ m) in 0.9% NaCl was made into the superficial layers of the SC. After 24–96 h, the eye contralateral to the injection site was treated with colchicine and 24 h later the retina was fixed, checked for transport of Fast Blue or microspheres and then processed for immunohistochemistry. A Leitz Orthoplan microscope with a Ploem fluorescent attachment and 'I2', 'N2' and 'A' filter blocks was used to detect FITC, TRITC and Fast Blue fluorescence respectively. The brain was cut and examined to confirm the location and extent of the injection site.



reported density distribution of ganglion cells in different parts of the retina: thus there are 1,000–1,400 immunoreactive cells mm^{-2} in the visual streak, compared with 125–200 cells mm^{-2} in the peripheral retina (Fig. 1b).

By nine days after optic nerve section, and at all subsequent survival times, there was a marked loss of SP-immunoreactive cells in the GCL and processes in the proximal inner plexiform layer in all retinal regions (Figs 1a, and 2c, d). The number of SP-immunoreactive cells in the GCL was reduced by approximately 90% in central retina and by 45–60% in peripheral retina (Fig. 1b). Most SP-containing cells remaining in the GCL had diameters of 8–10 μ m which is within the size range of the largest displaced amacrine cells^{7,15,18}. These retrograde degeneration studies clearly suggest that most SP-immunoreactive cells in the GCL are indeed ganglion cells. Therefore on the basis of these experiments and comparisons with ganglion cell counts from Nissl stained preparations^{16–18}, SP-immunoreactive cells probably constitute 25–35% of the ganglion cells in the GCL.

Following injection of either Fast Blue or rhodamine-labelled microspheres into the SC the distribution and density of retrogradely labelled ganglion cell bodies matched that reported previously⁷. In the central retina about 25% of Fast Blue labelled cells displayed SP-immunoreactivity (19–30% in several separate experiments) and 79% of the SP-immunoreactive cells in

the GCL were retrogradely labelled (Fig. 2e, f). Immunoreactive cells in the GCL lacking retrograde tracer in double-labelling experiments could be either displaced amacrine cells or ganglion cells that failed to transport detectable quantities of the tracer. These studies also support the conclusion that a substantial number of ganglion cells contain SP-immunoreactivity.

SP-immunoreactive ganglion cells project upon the contralateral dorsal lateral geniculate nucleus (dLGN), SC and accessory optic nuclei. In the dLGN a prominent plexus of immunoreactive axons and varicosities formed a narrow band some 120–250 μ m wide just below the optic tract (Fig. 3a, b). This plexus defines a concealed and previously undescribed lamina, which appears to correspond to the most superficial part of the outer layer of the alpha sector¹⁹. Immunoreactive axons, some of which could be traced into the dLGN from parent axons in the optic tract, gave rise to terminal arborizations characterized by spheroidal or sausage-shaped enlargements commonly 10 μ m in diameter and up to 30 μ m long (Fig. 3d). A similar immunoreactive plexus was present in the upper part of the stratum griseum superficiale of the SC and in the accessory optic nuclei. These plexuses were less prominent and disorganized between 3 and 6 days after, and absent 10 days or more after, contralateral optic nerve section or eye enucleation (Fig. 3a, c). However, such lesions had no detectable effect on the

Fig. 3 *a*, Frontal section of the diencephalon from a rabbit 16 days after unilateral optic nerve section showing immunoreactive fibres in the alpha sector of the ipsilateral dLGN just below the optic tract (OT), and loss of immunoreactive fibres from the contralateral dLGN. Note the plexuses of fine calibre immunoreactive processes in the intergeniculate leaflet (IGL) and the ventral lateral geniculate nucleus (vLGN) ($\times 5.6$). *b*, Enlargement of ipsilateral LGN from *a*, showing axon terminals and large varicosities in dLGN in a lamina just below the optic tract, and fine axons and varicosities in IGL and vLGN ($\times 11.2$). *c*, Enlargement of contralateral LGN from *a*, showing loss of immunoreactivity from the dLGN with preservation of the fine axonal plexuses in IGL and vLGN ($\times 11.2$). *d*, Immunoreactive axons and varicosities in the dLGN at higher magnification ($\times 280$).

Methods. For optic nerve section see Fig. 1 legend. For eye, enucleation, adult albino and Dutch belted rabbits ($N=9$) were anaesthetized with Sagatal (30–40 mg per kg), one eye irrigated with 1% amethocaine and the optic nerve exposed in the orbit through a lateral approach. The nerve and extraocular muscles were cut, the eyeball removed, the eyelids sutured and sprinkled with Cicatrin and 0.5 ml Intramycin given. Normal, eye-enucleated or optic-nerve-sectioned animals were anaesthetized, perfused with 0.9% NaCl in 0.1 M phosphate buffer followed by 4% paraformaldehyde (for light microscopy only) or with 4% paraformaldehyde and 0.1–0.5% glutaraldehyde followed by 4% paraformaldehyde (both containing 0.1 M lysine and 0.01 M sodium periodate) for light and electron microscopy. All perfusates were made up in 0.1 M phosphate buffer and were at room temperature. Brains were washed overnight in buffer (containing 20–30% sucrose in the case of brains to be cut on a freezing microtome) and sectioned in the frontal or horizontal plane at 25–30 μm (freezing microtome) or 50 μm (Vibroslice, Campden Instruments).

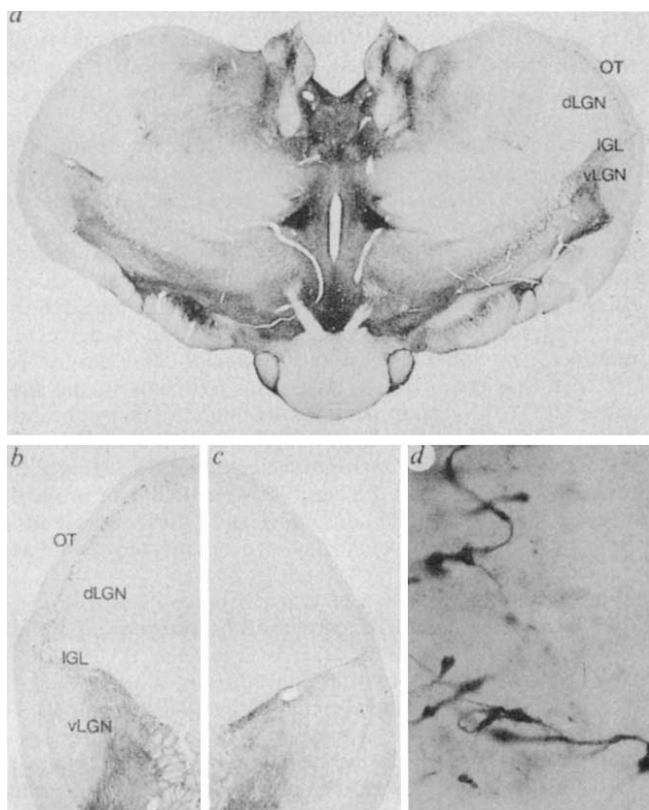
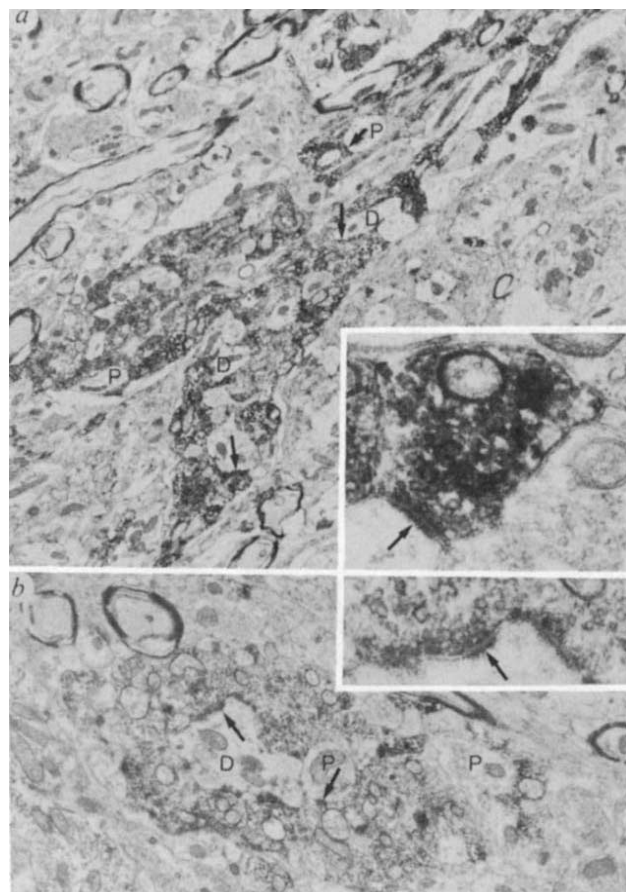


Fig. 4 *a, b*, Electron micrographs of synaptic glomeruli in the dLGN containing SP-immunoreactive retinal terminals and other non-immunoreactive profiles identified as dendrites or dendritic appendages of presumptive projection cells (D) or interneurons (P), which are postsynaptic to labelled profiles (arrows). The insets show synaptic contacts (arrows) between immunoreactive terminals and postsynaptic dendrites more clearly: the inset to *b* is an enlargement of part of the main panel; the inset to *a* is from a different glomerulus. Note the large, pale mitochondria, with swollen cristae, in the immunoreactive terminal profiles. Magnifications; in *a*, $\times 4,000$ (inset $\times 31,500$), in *b*, $\times 8,500$ (inset $\times 30,500$).

Methods. All tissue processed for electron microscopy was from rabbits perfused with a paraformaldehyde-glutaraldehyde mixture followed by paraformaldehyde alone and sectioned frontally at 50 μm on a Vibroslice (see legends to Figs 2 and 3). In some cases Triton X-100 was added to antibody-containing solutions to facilitate penetration. Reacted sections were osmicated (1% OsO_4 in 0.1 M phosphate buffer) stained in 1% aqueous uranyl acetate and embedded in Araldite. Thin sections were viewed in a Philips EM 301 with or without lead citrate staining.



corresponding immunoreactive plexuses in the ipsilateral dLGN (Fig. 3a, b) or SC, or on the plexuses of finer axons and smaller varicosities in the intergeniculate leaflet, ventral lateral geniculate, thalamic reticular or pretectal nuclei. These plexuses are therefore probably not of retinal origin (Fig. 3a-c).

Immunoreactive profiles in the dLGN were identified by electron microscopy as large axonal boutons in areas of complex neuropil (synaptic glomeruli) (Fig. 4a, b). The boutons were apposed to and commonly deeply invaginated by other, non-immunoreactive components of the glomeruli, which could be identified on the basis of previous ultrastructural studies of dLGN (for review see ref. 20) as the dendrites and dendritic appendages of presumptive projection cells and intrageniculate interneurons. Spherical synaptic vesicles and large, pale mitochondria, characteristic of retinal terminals in the mammalian dLGN²⁰, were apparent in the immunoreactive boutons, the synaptic relations of which were similar to those of identified retinal terminals in non-immunoreacted tissue. Specifically, they were presynaptic both to the dendritic appendages of projection cells and to the vesicle-containing dendrites and dendritic appendages of interneurons: they also established filamentous contacts with projection cell dendritic shafts.

These studies clearly suggest that substance P (and/or closely related tachykinins) may function as transmitters or modulators of rabbit retinal ganglion cells and provide the first clear demonstration of any putative transmitter or modulator in the axon terminals of the central visual pathway of a mammal. The presence of a prominent population of SP-containing ganglion cells in the rabbit retina, the recent suggestion that there are several peptide-containing ganglion cell types in the frog retina²¹ and reports of peptide bioactivity and immunoreactivity in optic nerve extracts^{1,22} raise the possibility that peptides are present in the retinal ganglion cells of other vertebrates including primates^{1,2,4}.

The functional characteristics of SP-containing ganglion cells are unknown and it is not yet clear how they relate to the subclasses of retinal ganglion cells described in other studies^{23,24}, although the distribution of their dendrites to the proximal inner plexiform layer suggests that they are likely to be involved in 'on-centre' pathways²⁵. Because SP-containing ganglion cells are distributed throughout the retina and their terminals in dLGN form an uninterrupted band from dorsal to ventral in the frontal plane and from rostral-lateral to caudo-medial in the horizontal plane we conclude that this lamina may contain a representation of the entire monocular portion of the contralateral visual hemifield. Finally, note that the immunoreactive plexus in the dLGN is partially overlapped by the terminal field of the projection to dLGN from the SC²⁶ and this region of dLGN may thus be comparable to the C-laminae in the cat, which also receive both retinal and collicular inputs²⁷.

We thank Ms M. Cilluffo and E. Franke for technical assistance, Mr P. Marx for help with the morphometric programmes, Drs L. Katz and B. Burge for rhodamine labelled microspheres and Drs C. Sternini and L. Kruger for helpful discussions. This work was partially supported by a grant from NIH and an Alfred P. Sloan Fellowship to N.B.

Received 20 February; accepted 13 March 1987.

- Brecha, N. in *Chemical Neuroanatomy* (ed. Emson, P. C.) 85-129 (Raven, New York, 1983).
- Brecha, N. C., Eldred, W., Kuljis, R. O. & Karten, H. J. *Prog. retinal Res.* 3, 185-226 (1984).
- Brecha, N., Hendrickson, A., Floren, I. & Karten, H. J. *Invest. Ophthalmol. Vis. Sci.* 23, 147-153 (1982).
- Höfelfelt, T. & Dahlstrom, A. Z. *Zellforsch. mikrosk. Anat.* 119, 460-482 (1971).
- Bentivoglio, M., Kuypers, H. G. J. M., Catsman-Berrevoets, C. E., Loewe, H. & Dann, O. *Neurosci. Lett.* 18, 25-30 (1980).
- Katz, L. C., Burkhalter, A. & Dreyer, W. J. *Nature* 310, 498-500 (1984).
- Vaney, D. I., Peichl, L., Wässle, H. & Illing, R.-B. *Brain Res.* 212, 447-453 (1981).
- Brecha, N. C., Oyster, C. W. & Takahashi, E. S. *Invest. Ophthalmol. Vis. Sci.* 25, 66-70 (1984).
- Parnavelas, J. G., Kelly, W., Franke, E. & Eckenstein, F. J. *Neurocytol.* 15, 329-336 (1986).
- Cuello, A. C., Galfré, G. & Milstein, C. *Proc. natn. Acad. Sci. U.S.A.* 76, 3532-3536 (1979).
- Harmar, A. J. *Trends Neurosci.* 7, 57-60 (1984).
- Famiglietti, E. V., Brecha, N. C. & Karten, H. J. *Neurosci. Abstr.* 6, 212 (1980).
- Cajal, S. R. y *La Célula* 9, 17-257 (1893).
- Hughes, A. & Vaney, D. I. *J. comp. Neurol.* 189, 169-189 (1980).

- Vaney, D. I., Peichl, L. & Boycott, B. B. *J. comp. Neurol.* 199, 373-391 (1981).
- Vaney, D. I. *J. comp. Neurol.* 189, 215-233 (1980).
- Oyster, C. W., Takahashi, E. S. & Hurst, D. C. *J. Neurosci.* 1, 1331-1346 (1981).
- Masland, R. H., Mills, J. W. & Hayden, S. A. *Proc. R. Soc. B* 223, 79-100 (1984).
- Holcombe, V. & Guillery, R. W. *J. comp. Neurol.* 225, 469-491 (1984).
- Jones, E. G. *The Thalamus* (Plenum, New York, 1985).
- Kuljis, R. O. & Karten, H. J. *J. comp. Neurol.* 240, 1-15 (1985).
- Unger, W. G., Butler, J. M., Cole, D. F., Bloom, S. R. & McGregor, G. P. *Expl Eye Res.* 32, 797-801 (1981).
- Amthor, F. R., Oyster, C. W. & Takahashi, E. *Proc. R. Soc. B* 217, 341-355 (1983).
- Vaney, D. I., Levick, W. R. & Thibos, L. N. *Expl Brain Res.* 44, 27-33 (1981).
- Bloomfield, S. A. & Miller, R. F. *J. Neurosci.* 6, 1-13 (1986).
- Holstege, G. & Collewyn, H. J. *comp. Neurol.* 209, 139-175 (1982).
- Torrealla, F., Partlow, G. D. & Guillery, R. W. *Neuroscience* 6, 1341-1360 (1981).

A novel class (H₃) of histamine receptors on perivascular nerve terminals

Shiro Ishikawa* & Nicholas Sperelakis

Department of Physiology and Biophysics, University of Cincinnati, College of Medicine, Cincinnati, Ohio 45267 0576, USA

Two types of histamine receptor, the H₁- and H₂-receptors^{1,2}, are found not only on vascular smooth muscle cells^{3,4} but on the perivascular autonomic nerve terminals. Activation of the prejunctional histamine receptors modifies transmitter release from the nerve terminals⁵⁻⁸. Recently, histamine was shown to inhibit its own release from depolarized slices of rat cerebral cortex. This phenomenon was found to be mediated by a novel class of histamine receptor, the H₃-receptor, that was pharmacologically distinct from the H₁- and H₂-receptors⁹. Up to now, there has been no indication whether this third class of histamine receptor is present in any tissue other than the brain. We report here that histamine depresses sympathetic neurotransmission in the guinea-pig mesenteric artery by interacting with histamine H₃-receptors on the perivascular nerve terminals. The pharmacological properties of these receptors are similar to those reported for the H₃-receptors in the brain. Our data provide evidence for the existence of H₃-receptors in the autonomic nervous system.

Under normal conditions, vascular smooth muscle (VSM) cells of the guinea-pig mesenteric artery were found to be electrically quiescent unless electrically stimulated and had a resting potential of -74.0 ± 0.3 mV ($n = 55$). Perivascular nerve stimulation with a brief impulse evoked an excitatory junction potential (e.j.p.) in the VSM cells^{10,11}. Repetitive stimulation (at a frequency of 1.0 Hz) produced a facilitation of the e.j.ps, the amplitude of which reached a steady level after 4 or 5 stimuli (Fig. 1, inset). These e.j.ps could be abolished by 3×10^{-7} M tetrodotoxin (TTX), consistent with their having a neurogenic origin.

Figure 1 shows that application of histamine (10^{-8} - 10^{-6} M) depressed the e.j.p. amplitude in a concentration-dependent manner, without modifying the passive membrane properties of the VSM cells (that is, there was no change in membrane potential or membrane resistance calculated from the current-voltage relationships). In the presence of 3×10^{-7} M TTX, 10^{-5} M histamine also was found to depolarize the VSM membrane; this depolarization was accompanied by an increase in membrane resistance. The histamine-induced depolarization was blocked by a selective H₂-receptor antagonist, cimetidine (3×10^{-6} M to 5×10^{-5} M), but not a selective H₁-receptor antagonist, chlorpheniramine (up to 10^{-6} M); these results indicate the presence of histamine H₂-receptors on the VSM cells. In the presence of 10^{-5} M cimetidine, 10^{-5} M histamine greatly depressed the e.j.p. amplitude to $18 \pm 3\%$ of the control value without affecting the passive membrane properties of the VSM cells. These data suggest that histamine depresses the e.j.p.

* Present address: Department of Pediatrics, Faculty of Medicine, Kyushu University, Fukuoka, Japan 812.

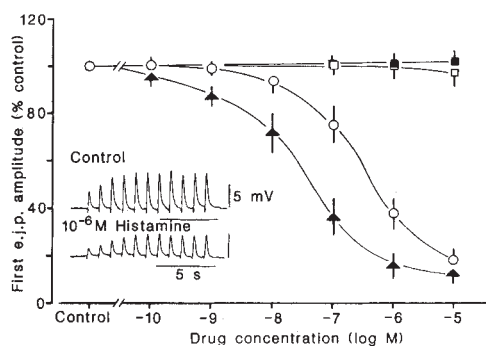


Fig. 1 Concentration-response curves showing the effects of four histamine receptor agonists on neurotransmission in the guinea pig mesenteric artery. Agonists as follows: $\circ$, histamine; $\blacksquare$, 2-methylhistamine; $\square$, dimaprit; $\blacktriangle$, N^{α} -methylhistamine. The amplitude of the first e.j.p. in a train of stimuli was expressed as a fraction of that in a histamine-free control. The e.j.p.s were recorded 5–10 min after application of each histamine receptor agonist. Vertical bars, s.e.m., $n=3-15$. Histamine-induced depression of the e.j.p. amplitude was not mimicked by a relatively selective H_1 -receptor agonist, 2-methylhistamine, or a selective H_2 -receptor agonist, dimaprit. A non-selective histamine derivative, N^{α} -methylhistamine, mimicked the histamine depression with tenfold higher potency than histamine itself (EC_{50} for histamine = 3.0×10^{-7} M and EC_{50} for N^{α} -methylhistamine = 3.0×10^{-8} M). These experiments were done in the absence of cimetidine. Inset, effects of histamine on the sympathetic neurotransmission. The train of e.j.p.s was evoked by repetitive nerve stimulation (30 μ s in duration, 25 V intensity, at 1 Hz for a total of 11 pulses). Histamine (10^{-6} M) inhibited the e.j.p. amplitude without modifying the passive membrane properties of the VSM cell. Both traces (above, control; below, histamine treated) are from the same VSM cell.

Methods. Guinea-pigs weighing 200–300 g, were stunned and bled. From the vascular bed of the jejunum, the mesenteric artery was dissected out (together with the parallel lymph and venous vessels and mesentery). Arteries with an external diameter of 120–200 μ m and about 5–10 mm long were mounted in a 2-ml organ bath for electrical recording. The electrical activity generated from a single smooth muscle cell was recorded using glass capillary microelectrodes filled with 3 M KCl having 50–80 M Ω resistance. A microelectrode was inserted into the VSM from the adventitial surface. Perivascular nerves were stimulated by drawing the proximal part of the artery into a suction electrode or by the point stimulation method¹¹. Current pulses of 30–100 μ s in duration and 20–50 V in intensity were supplied from an square-wave stimulator (Grass Model S88), and the resultant electrical activity was recorded on a rectifier (Gould, Model 2200S). The organ bath was superfused with Krebs solution (33–35 $^{\circ}$ C) at a flow rate of 2 ml min⁻¹. Normal Krebs bicarbonate solution contained (mM): Na⁺ 137.4, K⁺ 5.9, Ca²⁺ 2.5, Mg²⁺ 1.2, HCO₃⁻ 15.5, H₂PO₄⁻ 1.2, Cl⁻ 134.0, and glucose 11.5. This solution was bubbled with 95% O₂ and 5% CO₂, and the pH was 7.2–7.3. The following drugs were used: histamine-2 HCl, chlorpheniramine maleate, cimetidine, atropine sulphate, tetrodotoxin and yohimbine-HCl (Sigma); and 2-methylhistamine-2HCl, 2-pyridylethylamine-2HCl, 4-methylhistamine-2HCl, dimaprit-HCl, burimamide, impromidine-3HCl, metiamide, pyrilamine maleate and N^{α} -methylhistamine (Smith, Kline and French). The solutions containing the final concentrations of each drug were freshly prepared for each experiment. The values recorded were expressed as mean $\pm$ s.e.m. The statistical significance was determined using Student's *t*-test. A *P* value <0.05 was considered significant.

amplitude by interacting with a histamine receptor on the perivascular nerve terminals. Furthermore, it can be concluded that the inhibitory histamine receptors on the nerve terminals are more sensitive to histamine than are the H_2 -receptors on the VSM cells.

Various histamine receptor agonists were subsequently used to clarify which subclass of histamine receptor on the perivascular nerve terminals is involved in the depression of the e.j.p. amplitude (Fig. 1). A relatively selective H_1 -receptor agonist, 2-methylhistamine, did not modify the resting potential of

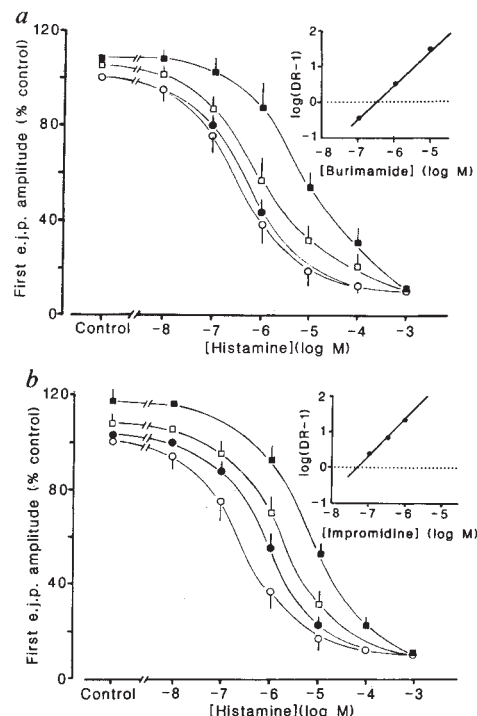


Fig. 2 Effects of burimamide and impromidine on the histamine depression of neurotransmission in the guinea-pig mesenteric artery. Burimamide (a) and impromidine (b) at a fixed concentration was used to antagonize increasing concentrations of exogenously applied histamine. Concentrations of antagonist denoted by symbols as follows: (in a) $\circ$, histamine alone; $\bullet$, 10^{-7} M; $\square$, 10^{-6} M; $\blacksquare$, 10^{-5} M burimamide; (in b) $\circ$, histamine alone; $\bullet$, 10^{-7} M; $\square$, 3×10^{-7} M; $\blacksquare$, 10^{-6} M impromidine. Responses are expressed relative to the first e.j.p. amplitude of histamine-free control (described in Fig. 1). Each point represents the mean $\pm$ s.e.m., $n=4-6$ for all groups. The insets represent the Schild plot of data, dose ratios (DR) being determined from $\frac{1}{4}$, $\frac{1}{2}$ and $\frac{3}{4}$ of maximal inhibitory concentrations of histamine (estimated graphically). If histamine and the antagonist compete for a common site, $\log(DR-1) = \log[Antagonist] - \log K_B$, where K_B = (apparent) dissociation constant for the antagonist-receptor interaction¹⁶; the data are plotted to examine this relationship. a, The slope for the regression of $\log(DR-1)$ versus $\log$ [burimamide] was 0.979 ± 0.03 , and the correlation coefficient was 0.986. The pA_2 value for burimamide was 6.48. b, The slope for the regression of $\log(DR-1)$ versus $\log$ [impromidine] was 1.11 ± 0.03 , and the correlation coefficient was 0.996. The pA_2 value for impromidine was 7.32. Effects of impromidine were investigated in the presence of 10^{-5} M cimetidine to prevent direct H_2 -receptor-mediated depolarization of the VSM cells.

the VSM cells or depress the e.j.p. amplitude even at 10^{-5} M. At a high concentration (10^{-5} M) a selective H_2 -receptor agonist, dimaprit, depolarized the VSM cells but failed to mimic the histamine depression of the e.j.p. amplitude. On the other hand, N^{α} -methylhistamine (10^{-8} M– 10^{-6} M), a non-selective histamine derivative, depressed the e.j.p. amplitude in a concentration-dependent manner without affecting the resting potential of the VSM cells; this agent was more potent than histamine itself (EC_{50} values for histamine and N^{α} -methylhistamine were 3.0×10^{-7} M and 3.0×10^{-8} M, respectively). These results indicate that histamine interacted with a receptor distinct from H_1 - and H_2 -subclasses to produce the depression of the e.j.p. amplitude.

To determine further which subclass of histamine receptor was responsible for the depression of the e.j.p. amplitude, we used several histamine receptor antagonists in an attempt to block the effect. The potent H_1 -receptor antagonists, chlorpheniramine and mepyramine (pyrilamine) (both up to 10^{-6} M) did not modify the resting potential of the VSM cell, nor did they affect the e.j.p. amplitude in the control or in the histamine-

Table 1 Effects of histamine receptor agonists and antagonists on the depression of e.j.p. amplitude

Agonist	Relative agonist potency (%)	Antagonist	Antagonist activity K_B (M)
Histamine	100	Chlorpheniramine	$>10^{-6}$ *
<i>N</i> $^{\alpha}$ -methylhistamine	1,000	Mepyramine	$>10^{-6}$ *
2-Methylhistamine	<1	Metiamide	$>3 \times 10^{-5}$ *
2-Pyridylethylamine	<1	Cimetidine	$>5 \times 10^{-5}$ *
4-Methylhistamine	<1	Burimamide	3.3×10^{-7}
Dimaprit	<1	Impromidine	4.8×10^{-8}

Histaminergic agonists (left columns) were investigated at different concentrations as described in Fig. 1. The relative potency was calculated as the ratio of EC_{50} for histamine to that of the tested agonist $\times 100$. The EC_{50} was defined as the concentration resulting in a half-maximal depression of the first e.j.p. amplitude. The EC_{50} for histamine was 3.0×10^{-7} M. Potential antagonists were applied in at least three different concentrations as described in Fig. 2. Values of K_B were calculated assuming a competitive antagonism, according to the Schild plot analysis.

* Compounds (up to the doses listed) with no effect on the histamine depression of e.j.p. amplitude.

depressed condition. The specific H_2 -receptor antagonists, cimetidine (up to 5×10^{-5} M) and metiamide (up to 3×10^{-5} M), also did not affect the resting potential or the e.j.p. amplitude. Simultaneous application of the H_1 - and H_2 -antagonists, clorpheniramine (10^{-6} M) and cimetidine (10^{-5} M), also failed to prevent the histamine depression. These results indicate that histamine interacted with a novel class of receptor to produce an inhibition of neuroeffector transmission.

Recent data suggest the presence of a novel subclass of histamine receptors, referred to as H_3 -receptors, in the central nervous system⁹. When endogenous histamine stores in the rat cerebral cortex were labelled by preincubation with tritiated histidine, [3 H]histamine release evoked by various depolarizing stimuli was progressively inhibited by the addition of exogenous histamine. This inhibitory effect was not mimicked by the selective H_1 - and H_2 -agonists, and was resistant to the selective H_1 - and H_2 -receptor antagonists. Further, the inhibitory effect was competitively inhibited by burimamide and impromidine, acting as putative H_3 -receptor antagonists. Therefore, we examined the effects of these compounds on the e.j.p. amplitude in the guinea-pig mesenteric artery.

Burimamide is less specific for the H_2 -receptors than other antagonists such as cimetidine and metiamide^{2,12}. In our experiments, burimamide, at concentrations over 10^{-6} M, slightly enhanced the e.j.p. amplitude under control conditions ($106 \pm 3\%$ of control at 10^{-5} M) without modifying the resting potential of the VSM cells. Moreover, this drug antagonized the histamine depression of the e.j.p. amplitude in a concentration-dependent manner. That the antagonism of histamine by burimamide could be surmounted was demonstrated by challenging fixed concentrations of antagonist with increasing concentrations of histamine. Figure 2a shows parallel shifts to the right of the concentration-response curve to histamine caused by burimamide, indicating competitive inhibition. Schild plot analysis of these data gave a straight line with a slope of 0.98 ± 0.06 , also compatible with burimamide acting as a competitive antagonist of histamine (Fig. 2a, inset). The negative logarithm of the affinity (pA_2) of burimamide was determined to be 6.48 ($K_B = 3.3 \times 10^{-7}$ M). This pA_2 value (6.48) is intermediate between that reported for a typical H_2 -receptor system, namely guinea-pig atrium and that for the H_3 -receptors postulated to be present in the rat cerebral cortex (5.11² and 7.52⁹, respectively). Table 1 summarizes the properties of histamine receptor agonists and antagonists on depression of e.j.p. amplitude.

Impromidine was reported to be a very potent but partial

histamine H_2 -receptor agonist¹⁰. Using the same protocol as for burimamide, the effects of impromidine on the histamine-induced depression were investigated in the presence of 10^{-5} M cimetidine to prevent direct H_2 -receptor-mediated depolarization of the VSM cells. In the presence of cimetidine, impromidine, at concentrations over 10^{-7} M, slightly enhanced the e.j.p. amplitude under control conditions ($117 \pm 5\%$ of control at 10^{-6} M) and antagonized the histamine-induced depression of e.j.p. amplitude. Figure 2b shows parallel shifts to the right of the concentration-response curve to histamine (maximal inhibitory effect was about 90% of control). Schild plot analysis gave a straight line with a slope of 1.11 ± 0.03 . The pA_2 value was 7.32 ($K_B = 4.8 \times 10^{-8}$ M), and closely resembled that obtained for impromidine in the rat cerebral cortex (7.52)⁹.

In this study, the putative H_3 -receptor antagonists, especially impromidine, enhanced the e.j.p. amplitude in a concentration-dependent manner in the absence of exogenous histamine (Fig. 2b, control). Similarly, the H_3 -receptor antagonists enhanced the K^+ -induced release of histamine in rat brain in the absence of exogenous histamine⁹. The mechanism for the stimulatory effect of impromidine on the e.j.p. amplitude is uncertain. Impromidine may possibly have been acting as an antagonist towards endogenous histamine released by neuronal activation^{13,14} and/or by basophilic granular cells around blood vessels¹⁵. Thus, the pharmacological properties of the prejunctional histamine receptor in guinea-pig mesenteric artery are similar to those reported for the H_3 -receptor of the brain.

However, a minor discrepancy remains. In the present experiments none of the concentrations of metiamide tested (10^{-6} , 10^{-5} and 3×10^{-5} M) affected the histamine-induced depression of e.j.p. amplitude. In contrast, concentrations of metiamide over 10^{-5} M partially prevented the inhibition of histamine release in the case of rat cerebral cortex⁹. It is possible that, in the guinea pig mesenteric artery, even higher concentrations of metiamide may be required to prevent histamine-induced depression of e.j.p. amplitude.

Our experiments, and the previously reported data for histamine action in the brain, suggest that H_3 -receptors are present at the prejunctional (or presynaptic) site of the autonomic nervous system and central nervous system. This class of histamine receptor is pharmacologically distinct from H_1 - and H_2 -receptors. We find the histamine H_3 -receptors to be very much more sensitive to histamine than are histamine H_2 -receptors on the VSM cells, so that they may have an important function in the physiological control of the circulation. These H_3 -receptors may produce vasodilation by inhibition of sympathetic tone.

We thank Dr J. Skidmore and Smith-Kline and French Research for many histaminergic compounds; Dr Y. Ishida and Dr J.-M. Arrang for advice and discussion; and Dr W. C. Cole for reading the manuscript and Miss A. Tolle for typing it. We also thank Drs H. Kuriyama, K. Ueda, H. Suzuki (Kyushu University) and N. Goya (Fukuoka Children Hospital) for their encouragement. These studies were supported by the NIH.

Received 22 December 1986; accepted 4 March 1987.

- Ash, A. S. F. & Schild, H. D. *Br. J. Pharmac.* **27**, 427-439 (1966).
- Black, J. W., Duncan, W. A. M., Durant, C. J., Ganellin, C. R. & Parsons, E. M. *Nature* **236**, 385-390 (1972).
- Chang, N. & Eyre, P. *Agents Actions* **5**, 277-295 (1975).
- Levi, R., Owen, D. A. & Trzeciakowski, J. in *Pharmacology of Histamine Receptors* (eds Ganellin, C. R. & Parsons, M. E.) 236-297 (Wright PSG, Bristol, 1982).
- Bevan, J. A., Duckles, S. P. & Lee, T. J-F. *Circulation Res.* **36**, 647-653 (1975).
- McGrath, M. A. & Shepherd, J. T. *Circulation Res.* **39**, 566-573 (1976).
- Powell, J. R. *J. Pharmac. exp. Ther.* **208**, 360-365 (1979).
- Suzuki, H. & Kou, K. *Pflügers Arch. ges. Physiol.* **399**, 46-53 (1983).
- Arrang, J.-M., Garbarg, M. & Schwartz, J.-C. *Nature* **302**, 832-837 (1983).
- Kuriyama, H. & Suzuki, H. *J. Physiol., Lond.* **317**, 383-396 (1981).
- Ishikawa, S. *Br. J. Pharmac.* **86**, 777-787 (1985).
- Ganellin, C. R. in *Pharmacology of Histamine Receptors* (eds Ganellin, C. R. & Parsons, M. E.) 10-102 (Wright PSG, Bristol, 1982).
- Tuttle, R. S. *Am. J. Physiol.* **213**, 620-624 (1967).
- Ryan, M. J. & Brody, M. J. *J. Pharmac. exp. Ther.* **174**, 123-132 (1970).
- Beaven, M. A. in *Pharmacology of Histamine Receptors* (eds Ganellin, C. R. & Parsons, M. E.) 103-145 (Wright PSG, Bristol, 1982).
- Waud, D. R. *Pharmac. Rev.* **20**, 49-88 (1968).

MHC antigens in urine as olfactory recognition cues

Prim B. Singh*, Richard E. Brown† & Bruce Roser*

* Immunology Department, AFRC Institute of Animal Physiology and Genetics Research, Babraham, Cambridge CB2 4AT, UK

† Subdepartment of Animal Behaviour, University of Cambridge, High Street, Madingley, Cambridge CB3 8AA, UK

The classical class I antigens of the major histocompatibility complex (MHC) are cell-surface glycoproteins which were originally discovered because they cause rapid rejection of cells or tissues grafted between unrelated individuals. These molecules are encoded by the *K*, *D* and *L* loci of the mouse MHC (and analogous loci in other species) which show extreme species polymorphism and a large number of alleles¹. In an outbreeding population 3.6×10^9 unique MHC class I phenotypes can be encoded by the 100 alleles at each of the *K* and *D* loci² and the 6 alleles at the *L* locus³. This level of polymorphism ensures that the cells and tissues of each unrelated individual are uniquely identified by their class I membrane-bound antigens. Like other membrane bound proteins, these class I molecules are anchored in the lipid bilayer by a hydrophobic domain encoded by exon 5. However, there have been reports of the occurrence of classical class I molecules in true solution in the blood of humans^{4,5}, mice^{6,7}, and rats⁸. We report here that classical polymorphic class I molecules in normal rats are constitutively excreted in the urine and that untrained rats can distinguish the smell of urine samples taken from normal donors that differ only at the class I MHC locus and therefore excrete different allelomorphs of class I molecules in their urine.

Pairs of non-competitive monoclonal rat alloantibodies chosen from the extensive panels of anti-class I antibodies produced by Diamond *et al.*⁹, were used in two-site (sandwich) enzyme-linked immunosorbent assays (ELISA) to detect molecules carrying two separate, allelic epitopes of the classical class I antigens, in serum, in freshly drawn, unclotted plasma (data not shown) and central lymph collected from the thoracic duct (Fig. 1a). Analysis of the molecules in lymph by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) revealed a heterodimeric molecule with a heavy chain of relative molecular mass 39,000 (M_r 39K) and a light chain of M_r 13K typical of classical class I antigens (Fig. 1b). Analysis of serum by gel filtration HPLC under non-dissociating conditions (Fig. 1c)

confirms that these molecules are true monomers and not membrane fragments or aggregates.

Sera of five strains of different MHC types DA(*RT1^{av1}*), PVG(*RT1^c*), F344(*RT1^{lo1}*), AS2(*RT1^f*) and BN(*RT1ⁿ*) were examined for the presence of soluble class I molecules using an inhibition binding assay, on DA red-cell targets, of ¹²⁵I-labelled mouse monoclonal antibody F16.4.4 (which detects a monomorphic determinant present on rat class I molecules of all strains¹⁰). This revealed the presence of soluble molecules in the sera of all the strains tested (data not shown). The amount of soluble antigen of donor type in tolerant rats (Fig. 1a) is roughly in proportion to the level of haemopoietic chimerism in these animals¹¹, which suggested that cells of the haemopoietic system may be a source of these molecules.

The concentration of class I molecules in solution in the serum was then estimated. The number of molecules on the erythrocyte surface was determined by Scatchard analysis¹² and, using erythrocytes as a calibrated source of antigen, the concentration of molecules on the serum was determined in an equilibrium competition assay by saturation analysis¹³, to be 350–380 ng ml⁻¹ (Table 1). The concentration in efferent lymph draining from lymphoid tissue is about twice this level (Fig. 1a). In a separate study the half-life of the serum molecule was measured and found to be 2.7 h (P.B.S., R.E.B. and B.J.R., manuscript in preparation).

The rapid loss of molecules from the blood suggested that they were either rapidly degraded or excreted. We therefore examined the urine of normal rats for the presence of these molecules using the two-site ELISA assay. They were readily detectable and calibration of the ELISA with the serum molecule showed that the concentration of class I molecules in the urine was quite variable and ranged between 40 and 190 ng ml⁻¹. Class I molecules purified from the urine by affinity chromatography on monoclonal antibody columns not only have the typical 39K heavy chain and 13K light chain associated with soluble class I molecules but also show a major protein band at 30K (Fig. 2a) suggesting that the class I molecules fragment in the urine.

Fractionation of class I antigen, which had been extracted from the urine, by gel filtration HPLC showed that antigenic material emerges from the HPLC column as a broad band of antigenic activity from about 50K down to 15K M_r (Fig. 2b). This is in sharp contrast to the behaviour of the class I molecules in serum, which emerge from the HPLC column as a well defined narrow peak between 68K and 43K (Fig. 1c). It therefore seems that once in the urine the majority of these molecules undergo rapid degradation to small fragments.

That polymorphic class I molecules of the A^{av1} and A^c type

Table 1 Quantitation of class I MHC molecules

MHC Class I antigen	Antibody	Scatchard analysis*		Saturation analysis†	
		<i>K</i> ($\times 10^8 M^{-1}$)	Molecules per RBC	<i>K</i> ($\times 10^8 M^{-1}$)	Serum concentration
A ^c	YR5/12	1.5	4,500 (PVG)	0.75	379 ng ml ⁻¹
A ^{av1}	JY1/116	1	10,300 (DA)	4.5	347 ng ml ⁻¹

RBC, red blood cell.

* Estimate of number of antigenic molecules on rat erythrocytes. The anti-A^{av1} antibody JY1/116⁹ and the anti-A^c antibody YR5/12³⁷ were iodinated using the lactoperoxidase method³⁸ at a ratio of 1 μ Ci to 1 μ g of purified IgG. These antibodies were then used to measure the number of antigenic sites on the relevant rat erythrocytes. In brief, titrated amounts of antibody were incubated with 25 μ l of a 10% (v/v) RBC suspension for 1 h at 4 °C and the cells separated from unbound antibody by centrifugation on a 200 μ l mixture of Di-n-butylphthalate (8 parts) and Di-nonyl phthalate (2 parts) (BDH). The tubes were frozen with solid CO₂ and then cut at the mid point of the frozen oil column. The RBC pellet and the floating supernatant were counted separately in a Packard model 5210 γ -counter (Packard) to give a precise measure of bound and unbound antibody. The antibody affinity and number of antigenic sites were then calculated by the method of Scatchard¹².

† Saturation analysis. One tenth of the above amount (25 μ l) of a 1% (v/v) RBC suspension was then used in an antibody-binding inhibition assay using 25 μ l of neat serum of the relevant type in 12 replicates. The three reactants, ¹²⁵I-labelled antibody, serum containing antigen, and RBCs were incubated for 1 h on a shaker at 4 °C to reach equilibrium. The cells were washed three times with PBS-azide + 2% (v/v) FCS and harvested onto glass fibre paper and washed again using a Skatron harvester (Flow Laboratories) and counted in a Packard γ -counter. Results were processed according to Ekins¹³. Serum in which the amount of class I molecule was known from saturation analysis was divided into aliquots and stored at -20 °C. This was always run in calibrated ELISA assays as a positive control. The amount of antigen in unknown samples run in these assays could then be estimated accurately by measurement of the horizontal displacement of regression lines drawn through the linear portions of the ELISA titration curves.

Fig. 1 Detection and characteristics of class I molecules in serum and lymph. *a*, ELISA titration of A^c antigen. Central lymph (▲) was collected from the thoracic duct of normal PVG rats by standard methods³⁴. Plasma, prepared by collecting aortic blood into chilled heparinized saline and immediate centrifugation at 4 °C, showed levels comparable to the serum (■). The level of donor PVG antigen in the serum of DA rats made neonatally tolerant of the (PVG × DA)F₁ hybrid strain (●) is low but readily detectable because of the constant, small background (broken line) in this assay (▼, control DA serum). *b*, SDS-PAGE analysis of membrane-bound and soluble class I molecules. Size markers, *M_r* in thousands. Membrane-bound A^{av1} molecules from DA liver have a heavy chain *M_r* of 47K and a β2-microglobulin light chain of 13K. The lymph molecule has a smaller heavy chain (*M_r* 39K). *c*, HPLC gel filtration of DA serum. DA serum (200 μl) was fractionated on an LKB TSK G3000SW column (LKB) running at 0.5 ml min⁻¹ in PBS-azide. The serum RT1.A^{av1} molecules, detected by ELISA, eluted as a sharp peak (shaded area) at the trailing edge of the serum protein peaks with an apparent *M_r* between that of BSA (68K) and ovalbumin (43K), showing that the molecules exist as soluble monomers in the circulation.

Methods. ELISA assays were carried out in rigid, non-sterile, flat-bottomed 96-well polystyrene plates (GIBCO). Pure monoclonal anti-class I antibody (200 μl of a 20 μg ml⁻¹ solution) in phosphate buffered saline containing 0.01% sodium azide (PBS-azide), pH 7.4, was used to coat individual wells. After blocking with PBS-azide containing 10% v/v fetal calf serum (FCS) for 1 h at 4 °C the plates were washed twice with PBS + 0.5% v/v Tween 20 (Sigma) (PBST). Aliquots (200 μl) of a solution containing the class I MHC antigen (or a titration thereof), were then added to the appropriate wells and incubated overnight at 4 °C. After six washes, 200 μl of a second non-competitive anti-class I monoclonal antibody coupled to biotin was added at 8 μg ml⁻¹ and incubated for 1 h. The plate was again washed six times before addition of 200 μl of streptavidin-coupled horseradish peroxidase (4 μg ml⁻¹) (Miles Laboratories) and incubated for twenty minutes at 4 °C. After four further washes, 200 μl of the substrate 3,3',5,5'-tetramethyl benzidine at 100 μg ml⁻¹ was added. The reaction was stopped after 5 min with 50 μl of 2 M H₂SO₄ and the plates read in a Titertek Multiskan ELISA plate reader (Flow Laboratories). In the assay for detection of the A^{av1} class I MHC antigens of the DA strain the first antibody on the plate was JY3/109 (haplotype specificity, *av1*+, *c*-, *f*-, *h*-, *l*-, *n*-, *o*+, *u*-, *b*+), whereas the second biotin-labelled antibody was JY1/116 (ref. 9) (haplotype specificity, *av1*+, *c*-, *f*-, *h*-, *l*-, *n*-, *o*-, *u*-, *b*+). For the detection of the A^c antigens of the PVG strain, the non-competitive antibodies were YR5/310 (haplotype specificity, *av1*-, *c*+, *f*-, *h*+, *l*-, *n*-, *o*-, *u*-), on the plate, and YR5/12²⁸ (haplotype specificity, *av1*-, *c*+, *f*-, *k*-, *l*-, *n*+, *o*-, *u*-) as the biotin-labelled second stage. All are rat IgG alloantibodies. Rats were tolerized and tolerance authenticated as in ref. 13. In *b*, DA strain rat liver membranes solubilized in 1% *n*-octylglucopyranoside (Sigma) were purified by affinity chromatography³⁵. The JY1/116 anti-RT1.A^{av1} affinity column was washed exhaustively before elution with 0.015 M carbonate buffer pH 11.0 containing 0.5% (w/v) *n*-octylglucopyranoside. The eluted fractions were concentrated using a CX30 concentrator (Millipore) before running on SDS-PAGE³⁶ and staining with Coomassie blue. To purify the soluble RT1.A^{av1} molecules from lymph, it was spun at 1.05 × 10⁵ g for 1 h so that all the lipid and lipoproteins floated and could be removed. The opalescent lymph was loaded onto the columns and washed as above except there was no detergent in the buffers. RT1.A^{av1} was eluted with 0.5% (w/v) *n*-octylglucopyranoside in 0.015 M carbonate buffer pH 11.0, concentrated and run on SDS-PAGE as for the solubilized membrane material.

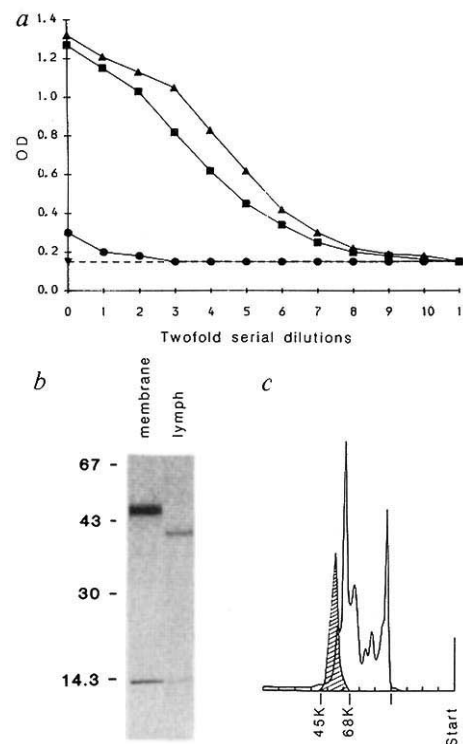


Fig. 2 Detection and characteristics of soluble class I molecules in urine. *a*, SDS-PAGE analysis of A^{av1} molecule in urine. The A^{av1} molecule from 500 ml of DA urine was purified by affinity chromatography as described for lymph in Fig. 1 legend. Note that the *M_r* of the class I heavy chain in urine is 39K as it is in the serum. The 13K light chain of β2-microglobulin is also seen indicating that the molecules detected by ELISA are typical class I molecules. In addition there is a major band at 30K which is not seen in either the lymph or membrane form of this molecule. *b*, HPLC gel-filtration of urine class I molecules. Class I molecules purified from the urine as above were run in HPLC and the fractions assayed as in Fig. 1c. Note that antigenic material elutes from this column with a wide range of apparent *M_r* from 50K down to about 15K.

and their degradation products in the urine of normal rats could be detected by ELISA prompted us to see whether rats could also detect urinary odours associated with these molecules. In these experiments we used the habituation, dishabituation method¹⁴ in which untrained detector animals were habituated to the odour of urine samples from donors of one strain and then tested for dishabituation when exposed to urine from donors of the other strain. To exclude a function for any variable other than the MHC class I molecules in these experiments, we used throughout, male rats from the PVG congenic series¹⁵. PVG-RT1^u detector animals were tested for their ability to discriminate between the urines of PVG donors and PVG.R1 donors. The latter are congenic recombinant animals and differ from PVG only at the classical class I genetic locus RT1.A¹⁵. The results (Fig. 3) show that the PVG-RT1^u rats readily discriminate the odour of PVG urine containing the A^c molecule from PVG.R1 urine containing the A^{av1} molecule. When this

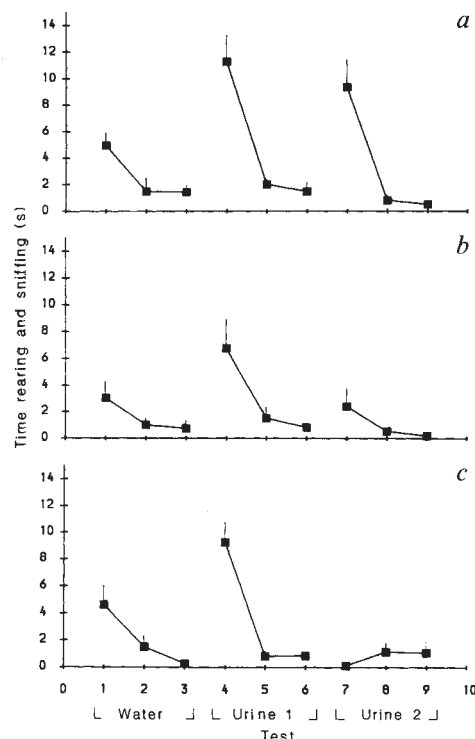
experiment was repeated with urine presentations in the reverse order (PVG.R1 urine followed by PVG urine) the results were identical (data not shown).

The use of urine to mark the environment is a well-known behavioural trait in many species. On the basis of the odour of a urine sample, animals can identify a wide range of phenotypic characters including the species, strain and sex of the donor^{16,17}. The precision of this olfactory discrimination extends to the ability of mice and rats to identify MHC disparate animals on the basis of urinary odour¹⁸ (Fig. 3). These cues are used in mate selection¹⁹ and can also control fertility²⁰. Thus up to 90% of pregnant female mice abort their pre-implantation embryos when exposed to the odour of urine from a foreign male²¹ even when it is from an inbred strain congenic with that of the original stud and differs only at the MHC²² or is a mutant (*bm1*) at one class I MHC locus²³. The only known difference between *bm1* and the C57BL/6 wild type are point mutations in the

Fig. 3 Habituation-dishabituation tests. *a*, The urine samples were from different strains and, after the 3 tests with water, the first (PVG) urine sample was investigated significantly more than the third water sample. The response habituated rapidly and when the second (PVG.R1) urine sample was presented, the subject dishabituated, showing a mean time of 9.2 s investigating the odour. Response to the second odour again habituated rapidly. *b*, Urine samples from two individual males of the same strain (PVG-1, PVG-2) were presented. There was no significant dishabituation. *c*, Two urine samples from the same individual (PVG-3, PVG-3) were presented and there was no evidence of dishabituation. Statistical results were as follows. Randomized block (repeated measures) analyses of variance were done separately for each experiment. Post-hoc analyses were done using Newman-Keuls tests³². In *a* there were significant differences among the 9 means ($F = 24.249$; d.f. = 8, 104; $P < 0.001$) and post-hoc tests showed that significantly more time was spent investigating odours on tests 4 and 7 than on any other test ($P < 0.01$).

In *b* there were significant differences among the 9 means over all 8 subjects ($F = 17.730$; d.f. = 8, 120; $P < 0.001$), more time was spent sniffing the odour on test 4 than on any other test ($P < 0.01$). There was no significant difference between tests 6 and 7. In *c*, there were significant differences among the 9 means for all 8 subjects ($F = 21.053$; d.f. = 8, 120; $P < 0.001$), and the time sniffing on test 4 was significantly more than on all other tests ($P < 0.01$). Time sniffing on test 7 did not differ from that on test 6.

Methods. Urine was collected from 10 individual PVG and 10 individual PVG.R1 males the same age as the subjects and frozen in aliquots until used. Each subject was given two urine samples from known individuals. Eight adult (90–100 day-old) male PVG-RT1^u rats were housed in groups of two for two weeks in a reversed 12:12 h light-dark cycle and handled each week. They were then given a 15 min habituation to the test area (an opaque plastic cage measuring 29.6 × 23.6 × 14.6 cm with a half-inch square wire mesh top which rose another 5.6 cm above the cage rim). Each rat was then tested once per week in a habituation-dishabituation test in which 9 sequential 2-min odour presentations were given. For these presentations, 0.1 ml of liquid was placed on a 7-cm diameter disk of Whatman No. 1 filter paper and this was taped to the mesh top of the cage, with the centre of the disk 13.5 cm from the floor of the arena. The odorized disk was replaced every 2 min, by removing the top of the cage and replacing it with another top carrying a new odorized disk. Tops and arenas were washed in 70% ethanol after each test and filter papers were used only once. On the first three 2-min tests, water was placed on the filter paper, so the subject had a 6-min habituation to the test protocol before the first urine sample was presented for the next three 2-min trials. On trial 7, the second urine sample was presented for three trials. As described by Sundberg *et al.*¹⁴, when the subject realizes there is an odour present, it lifts its head, walks or turns toward the odour source while sniffing the air and moving its head back and forth and sideways and finally rears and sniffs the filter paper disk. The time spent rearing on the hind legs and sniffing, with the nose within 1 cm of the disk throughout each 2-min period, was recorded by an observer using a stopwatch. Each of the eight rats was tested in three habituation-dishabituation tests. The observer was 'blind' to the order of testing odourant samples which were prepared by an assistant according to a random list.



class I genes resulting in MHC molecules which differ by 3 amino acids. This implies that the olfactory cues in the urine are not derived from genes closely linked to the class I locus but directly or indirectly from the product of the class I genes themselves.

Class I molecules are large glycoproteins and essentially non volatile, so it is difficult to see how they can be detected by smell. Although there are mechanisms described for the detection of²⁴ and behavioural responses to²⁵ non-volatile urinary molecules by means of the vomeronasal organ, there is clear evidence that volatile odours are used by rodents to identify the MHC type of urine donors²⁶. There are several possible explanations for this paradox: Firstly, the odoriferous components of the urine could be small volatile fragments of the MHC molecules themselves. This mixture of small fragments, which would of course be unique to the particular polymorphic class I molecules excreted, may be the specific odorant and it is of interest that rats which can discriminate the PVG and PVG.R1 strains by smelling their urine which contains the A^c and A^{av1} molecules in both degraded and intact forms (Fig. 2), fail to distinguish between these strains by the smell of their serum³³ which contains the molecules only in the intact form (Fig. 1c). This is also true in the mouse (E. A. Boyse, personal communication).

Secondly, MHC glycoproteins are well known as associative molecules. Thus the normal immune response is triggered, not by macromolecular antigens themselves, but by small processed fragments of the antigens which are associated with the MHC class I molecules of the host on the surface of antigen presenting cells. The T-lymphocyte receptor recognises only this associative complex and is triggered by it^{27,28}. This associative recognition is specific to allelic epitopes on MHC molecules and is the mechanism of MHC restriction of cellular interactions in the immune response^{29,30}.

We propose that the ability of MHC class I molecules to associate in a selective way with other small molecules could also be the mechanism by which a unique mixture of volatile, endogenous metabolites is transported by class I MHC glycoproteins from the blood into the urine. This mixture, although derived from a similar metabolic pool in each individual, would be unique to the particular class I molecules involved in its transport and would therefore impart an individual-specific odour to the urine. The volatile molecules we postulate have not yet been identified but strong candidates may be the small volatile components identified by Schwende *et al.* in unique mixture patterns in the urine of congenic mice³¹.

Whatever the mechanism of detection (direct identification of MHC glycoproteins by the vomeronasal organ or identification of volatile components by olfaction) the fact that the classical class I transplantation antigens of the MHC are constitutively excreted in relatively large amounts in the urine of normal animals indicates a second major biological role for these glycoproteins. Thus, they not only exert control (in their membrane-bound form) over the specificity of cellular interactions in the immune response by virtue of their function as restricting elements, but they also may exert control (in their soluble form) over the specificity of interaction between individuals in a species by virtue of their excretion in the urine and detection by olfaction.

This work was supported by grants from the MRC, and the National Science and Engineering Research Council of Canada. P.B.S. was supported by an AFRC studentship and a Magdalene College Bye-fellowship. We thank E. A. Boyse, G. W. Butcher and J. C. Howard for comments on the manuscript, E. C. Bate-Smith and R. van Aarde for useful discussions, and A. Diamond, J. C. Howard and G. W. Butcher for monoclonal antibodies and advice on their use.

Received 10 February; accepted 24 March 1987.

1. Klein, J. *The Natural History of the Histocompatibility Complex* (Wiley, New York, 1986).
2. Duncan, W. R., Wakeland, E. K. & Klein, J. *Immunogenetics* **9**, 261-272 (1979).
3. Klein, J. in *Histocompatibility Antigens: Structure and Function. Receptors and Recognition, Series B* (eds Parham, P. & Strominger, J.) 221-239 (Chapman and Hall, London, 1982).
4. Van Rood, J. J., Van Leeuwen, A. & Van Santen, M. C. T. *Nature* **226**, 366-367 (1970).
5. Vincent, C. & Revillard, J. P. *Transplant Proc.* **11**, 1301-1302 (1979).
6. Callahan, G. N., Ferrone, S., Allison, J. P. & Reisfeld, R. A. *Transplantation* **20**, 431-433 (1975).
7. Callahan, G. N., Ferrone, S., Penlik, M. D., Reisfeld, R. A. & Klein, J. *J. Immun.* **117**, 1351-1355 (1976).
8. Zimmerman, F. A. et al. *Transplant. Proc.* **11**, 571-577 (1979).
9. Diamond, A. D. et al. *Eur. J. Immun.* **14**, 405-412 (1984).
10. Hart, D. N. J. & Fabre, J. W. *Transplant. Proc.* **13**, 1329-1332 (1981).
11. Roser, B. & Dorsch, S. E. *Immunol. Rev.* **46**, 55-86 (1979).
12. Scatchard, G. *Ann. N.Y. Acad. Sci.* **51**, 660-672 (1949).
13. Ekins, R. P. *Br. med. Bull.* **30**, 3-11 (1974).
14. Sundberg, H., Doving, K., Novikov, S. & Ursin, H. *Behavioural and Neural Biology* **34**, 113-119 (1982).
15. Butcher, G. W. & Howard, J. C. *Handbook of Experimental Immunology*, 4th edn (ed. Weir, D. M.) 101.1-101.18 (Blackwell, Oxford, 1986).
16. Brown, R. E. & Macdonald, D. W. (eds) *Social Odours in Mammals* (Oxford University Press, 1985).
17. Brown, R. E. *Advances in the Study of Behaviour* **10**, 103-162 (1979).
18. Beauchamp, G. K., Yamazaki, K. & Boyse, E. A. *Sci. Am.* **253**, 66-72 (1985).
19. Yamazaki, K. et al. *J. exp. Med.* **144**, 1324-1335 (1976).
20. Bruce, H. M. *Nature* **1284**, 105 (1959).
21. Dominic, C. J. *J. Reprod. Fert.* **8**, 266-267 (1964).
22. Yamazaki, K. et al. *Science* **221**, 186-187 (1983).
23. Yamazaki, K. et al. *Proc. natn. Acad. Sci. U.S.A.* **83**, 740-741 (1986).
24. Wysocki, C. J., Wellington, J. L. & Beauchamp, G. K. *Science* **207**, 781-782 (1980).
25. Johns, M. A., Feder, H. H., Komisaruk, B. R. & Mayer, A. D. *Nature* **272**, 446-447 (1978).
26. Yamazaki, K. et al. *J. exp. Med.* **150**, 755-760 (1979).
27. Townsend, A. R. M., Bastin, J., Gould, K. & Brownlee, G. G. *Nature* **324**, 575-577 (1986).
28. Maryanski, J. L., Pola, P., Corradin, G., Jordan, B. R. & Cerottini, J.-C. *Nature* **324**, 578-579 (1986).
29. Bevan, M. J. & Fink, P. J. *Immun. Rev.* **42**, 3-19 (1978).
30. Zinkernagel, R. M. & Doherty, P. C. *Adv. Immun.* **27**, 51-177 (1979).
31. Schwende, F. J., Jorgenson, J. W. & Novotny, M. *J. chem. Ecol.* **10**, 1603-1615 (1984).
32. Kirk, R. E. *Experimental Design: Procedures for the Behavioural Sciences* (Brooks/Cole, Belmont, California, 1968).
33. Brown, R. E., Singh, P. B. & Roser, B. J. *Physiology and Behaviour* (Pergamon, Oxford, in the press).
34. Ford, W. L. in *Handbook of Experimental Immunology* (ed. Weir, D. M.) 23.1-23.22 (Blackwell, Oxford, 1978).
35. Parham, P. *J. biol. Chem.* **254**, 8709-8712 (1979).
36. Laemmli, V. K. *Nature* **227**, 680-685 (1970).
37. Stephenson, S. P., Livingstone, A. M., Butcher, G. W. & Howard, J. C. *Transplant Proc.* **17**, 1844-1846 (1985).
38. Heusser, C., Boseman, M., Nordin, J. H. & Isliker, H. *J. Immun.* **110**, 820-828 (1973).

Independence of *fushi tarazu* expression with respect to cellular density in *Drosophila* embryos

William Sullivan

Department of Biochemistry and Biophysics,
University of California, San Francisco, San Francisco,
California 94143-0448, USA

Regularly spaced repeated morphological structures are a common developmental theme among higher eukaryotes. In *Drosophila*, this is evident in the repeated segments of the larval and the adult cuticle. It has been demonstrated through cell transplantation¹ and more recently through molecular techniques² that these repeated segmental units are established as early as nuclear cycle 14 in the blastoderm embryo. A number of genes have been shown to express their transcripts², and in two instances their protein products^{3,4}, in a spatially restricted manner at this early stage. Immunofluorescence probes against the protein product of one such gene, *fushi tarazu* (*ftz*), reveal that it is distributed in seven evenly spaced stripes across the cellularized cycle-14 blastoderm embryo³. The mechanisms that determine such spatial patterns of gene expression are of fundamental importance for the development of multicellular organisms, but in no case are they well understood. Here we examine the *ftz* pattern on blastoderm embryos derived from *maternal-haploid 1182* (*mh 1182*)^{5,6} and *daughterless-abo-like* (*dal*)⁷ females which possess cell densities and sizes both above and below the wild-type levels. The number, spacing and width of the *ftz* protein positive bands are not altered in these abnormal embryos relative to the wild-type pattern, suggesting that the mechanism by which distance is measured with respect to the *ftz* protein is independent of cell size and density.

After 13 synchronous nuclear divisions without accompanying cytokinesis, the *Drosophila* embryo reaches cycle 14 as a syncytium of some 6,000 nuclei that are arranged in a uniform monolayer over the surface of the embryo. During interphase of this cycle, cellularization occurs⁸. Nuclear density counts for wild-type embryos indicate there are 65-80 cells per 2,200 μm^2 at this stage with an average cell diameter of 5.2-5.8 μm .

Females homozygous for the *mh1182* allele produce embryos in which the nuclei are initially haploid. As a result, these nuclei undergo an extra division before cellularizing⁹. Therefore at cellular blastoderm, embryos possess smaller cells at about twice the density (130-180 cells per 2,200 μm^2) of that found in wild-type embryos (Fig. 1).

Embryos derived from *dal* females often possess cellular blastoderm densities both above and below that found in normal

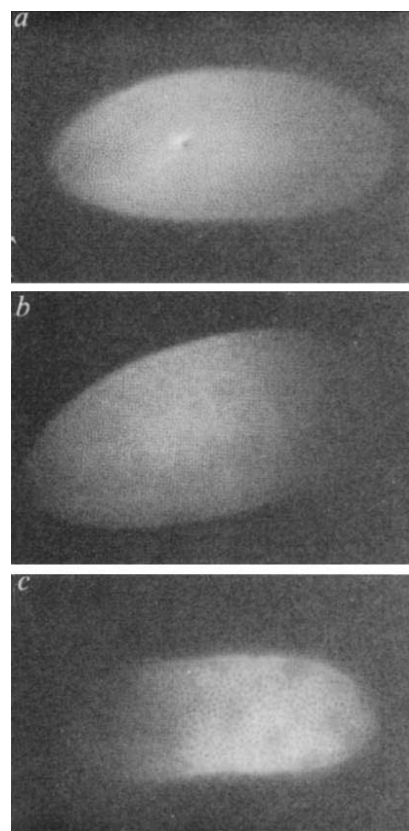


Fig. 1 Photographs of *dal*-derived (*a* and *c*)⁷ and *mh1182* (*b*)^{5,6} embryos at cellular blastoderm. These embryos have nuclear densities of *a*, 99 per 2,200 μm^2 ; *b*, 173 per 2,200 μm^2 and *c*, 48 per 2,200 μm^2 . They were stained by treatment with a rabbit α -phosphotyrosine membrane-specific antibody that was visualized with a rhodamine-labelled secondary antibody^{8,23}. Comparisons of these photographs with the corresponding DAPI-stained images demonstrates that the nuclear density counts correspond well with cellular densities and sizes. The α -phosphotyrosine antibody was a gift of J. Wang.

embryos. Those with cellular densities below the wild-type levels (less than 50 cells per 2,200 μm^2) possess larger possibly polyploid cells (Fig. 1). In addition, a small fraction of the *dal*-derived embryos are found to contain large patches of cells differing in size and density from those in the rest of the embryo (Fig. 2). The primary defect which results in these abnormal cell densities at cellular blastoderm has not yet been determined.

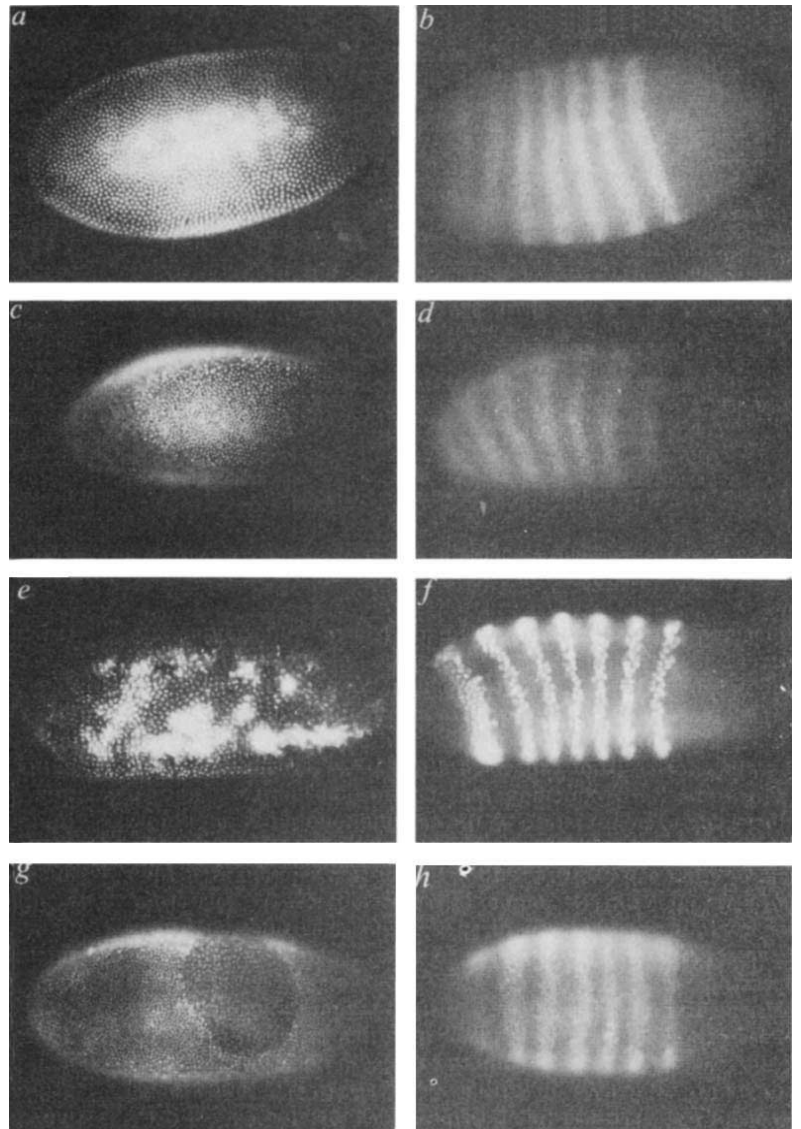


Fig. 2 Photographs of three embryos double stained with DAPI (left-hand panels) and the antibody against the *ftz* protein (right-hand panels). *a, b*, Staining pattern of a wild-type embryo (cellular density = 72 and cells per stripe width = 3.3); *c, d*, staining pattern on an embryo with increased cellular density (cellular density = 173 and cells per stripe width = 5.7); *e, f*, staining pattern on an embryo with a decreased cellular density (cellular density = 49 and cells per stripe width = 1.7); *g, h*, a *dal*-derived embryo with a patch of cells at one-third the cellular density of that of the surrounding cells (52 per 2,200 μm^2 compared to 141 per 2,200 μm^2). The immunofluorescence staining procedure was as described³, using an antibody against *ftz* protein that was a gift of S. B. Carroll and M. P. Scott.

Here the cellular densities at blastoderm were determined using the DNA-specific stain DAPI (diamidino phenyl indole). To rule out the possibility that many of the nuclei in mutant embryos are excluded from the cellularization process, wild-type and mutant embryos were double stained with DAPI and a membrane-specific antibody¹⁰. The membrane staining shown in Fig. 1 demonstrates that the nuclear density counts obtained by DAPI staining accurately reflect cell size and density.

At the cellular blastoderm stage in wild-type embryos, the *ftz* protein is expressed in seven regularly spaced stripes³. The six anterior-most stripes average from three to four cells in width and the seventh is about five cells wide. The intervals between the stripes also average three to four cells in width³.

Measurements on six wild-type cellular blastoderm embryos gave densities of between 68 and 77 nuclei per 2,200 μm^2 . The average width of the six most anterior *ftz*-positive stripes in these embryos was $4.2 \pm 0.6\%$ of the total egg length ($N = 36$). The average interstripe width was $3.8 \pm 0.8\%$ of the total egg length ($N = 36$). The larger seventh stripe averaged $6.6 \pm 0.4\%$ of the total egg length ($N = 6$). Table 1 provides a more detailed description of these data. For five wild-type embryos ranging in nuclear density from 68 nuclei to 75 nuclei per 2,200 μm^2 the average number of *ftz*-positive cells per stripe width (the three anterior-most stripes were scored in each embryo) ranged from 3.1 to 3.7 (Table 1).

Females homozygous for the maternal-effect mutation *dal* produce a class of cellular blastoderm embryos with larger cells at a lower density compared with wild-type embryos. Measurements on eleven such embryos revealed densities ranging from 25 to 57 cells per 2,200 μm^2 (Table 1). The width of the six anterior-most stripes and interstripes averaged $3.8 \pm 0.6\%$ ($N = 66$) and $4.0 \pm 1.0\%$ ($N = 66$) of the total egg length respectively. The width of the seventh stripe was $6.8 \pm 0.6\%$ ($N = 11$). These values are similar to the wild-type values (Table 1), but the number cells per stripe width ranges from 1 to 2 rather than the 3 to 4 found for wild-type embryos. For seven embryos ranging in nuclear density from 38 to 51 nuclei per 2,200 μm^2 , the average number of *ftz* cells per stripe width (the three anterior-most stripes were scored) ranged from 1.4 to 1.7 (Table 1).

Females homozygous for either *dal* or *mh1182* produce a fraction of embryos with smaller cells at higher densities in the cellular blastoderm embryos. Measurements on eight of these embryos indicated densities between 130 and 174 cells per 2,200 μm^2 , about twice that found in the wild-type embryos (Table 1). The average width of the six anterior-most stripes and interstripes were $4.4 \pm 1.0\%$ ($N = 40$) and $3.8 \pm 0.8\%$ ($N = 40$) of the total egg length respectively. The seventh stripe averaged $5.2 \pm 0.6\%$ ($N = 8$). Again these values are similar to those found in the wild-type embryos. However, in these embryos the number of cells spanning the width of the stripe

Table 1 Stripe and interstripe widths and number of cells per stripe width in embryos of differing cellular densities

	Cellular density (X)		
	$X < 60$	$68 < X < 77$	$X > 138$
1st stripe	0.021 ± 0.002	0.022 ± 0.004	0.026 ± 0.003
1st inter	0.025 ± 0.004	0.022 ± 0.003	0.020 ± 0.003
2nd stripe	0.018 ± 0.002	0.022 ± 0.003	0.027 ± 0.004
2nd inter	0.019 ± 0.004	0.019 ± 0.004	0.016 ± 0.003
3rd stripe	0.018 ± 0.003	0.020 ± 0.003	0.017 ± 0.007
3rd inter	0.019 ± 0.005	0.017 ± 0.003	0.015 ± 0.003
4th stripe	0.016 ± 0.003	0.020 ± 0.003	0.020 ± 0.003
4th inter	0.016 ± 0.003	0.018 ± 0.003	0.016 ± 0.003
5th stripe	0.018 ± 0.003	0.020 ± 0.004	0.018 ± 0.002
5th inter	0.019 ± 0.002	0.022 ± 0.002	0.013 ± 0.002
6th stripe	0.021 ± 0.004	0.022 ± 0.002	0.022 ± 0.004
6th inter	0.024 ± 0.002	0.021 ± 0.003	0.019 ± 0.003
7th stripe	0.034 ± 0.003	0.033 ± 0.002	0.026 ± 0.003
Average of stripes 1-6	0.019 ± 0.002	0.021 ± 0.003	0.022 ± 0.005
Average of interstripes 1-6	0.020 ± 0.005	0.019 ± 0.004	0.017 ± 0.004
Number of cells per stripe width	1 to 2	3 to 4	4 to 6
Number of eggs scored	11	6	8

A summary of the average stripe and interstripe (inter) widths (in mm) for the seven stripes (1 = anterior-most) in three sets of embryos grouped according to their cellular density. The average number of cells per stripe width is also indicated. The embryos have all been standardized to a total length of 0.5 mm. To determine the average stripe width and the number of cells spanning a stripe, photographs of both the DAPI-stained image and the *ftz*-stained image were taken at the same position and focal plane. Projection and tracing of the resulting images enabled those two images to be superimposed. From the *ftz*-stained image, a digitizer was used to determine the area and length of each stripe and interstripe, from which average stripe and interstripe widths were determined. This method was used because the stripes are often irregularly shaped along their borders. The width was then standardized to an embryo with a total length of 0.5 mm. The average number of cells spanning a stripe was determined using the tracings of the superimposed DAPI-stained image. The number of rows of nuclei in each stripe was determined and then the total number of nuclei in that area was counted. Dividing the total number of nuclei by the number of rows results in an estimate of the average number of cells spanning a stripe.

embryos. For six embryos ranging in nuclear density from 137 to 174, the average number of *ftz*-positive cells per stripe width (the three anterior-most stripes were scored) ranged from 4.0 to 7.4.

A small fraction of the embryos derived from *dal* females are mosaic in cell density and size at blastoderm. The width of the *ftz* positive stripes are not altered as they cross patches of cells varying in cell size and density from those in the rest of the embryo (Fig. 2 and Table 2). The previous results show that the stripe width does not change with cell density, but the number of cells spanning the width of the stripe varies proportionately with cell density (and therefore inversely with cell size). In the mosaic embryos, this phenomenon is observed even within individual stripes. In addition, even at the well-defined boundary between the groups of cells with different densities and sizes, the width of the *ftz* stripe is not appreciably altered (Fig. 2).

The approach for this study was motivated from the earlier work of Held¹¹ and Santamaria¹² in which they showed that the spacing of bristles increases from haploid to diploid to triploid patches of cells in adults mosaic for haploid and triploid tissue. Because cell size increases with increasing ploidy¹³, Held suggested that the interval between bristles must be measured as a fixed number of cells or as a distance which depends on cell diameter. Because pattern elements which are specified by gradients would not be expected to be affected by cell size and

Table 2 Average width and number of cells per stripe in stripes which span regions mosaic in nuclear density

Embryo designation	1		2		3	
	Cellular density (cells per 2,200 μm^2)					
	62	120	52	141	50	138
1st stripe	—	0.024	0.019	0.019	0.025	0.022
1st inter	—	0.030	0.024	0.026	0.017	0.020
2nd stripe	0.029	0.027	0.021	0.021	0.017	0.022
2nd inter	0.017	0.012	0.021	0.021	—	0.017
3rd stripe	0.020	0.021	0.021	0.017	—	0.016
Average no. of cells per stripe width	2.8	5.6	2.4	4.3	2.0	3.9

Measurements (in mm) of the average stripe and interstripe widths (inter) and the number of cells spanning a stripe in three embryos that were mosaic in their cellular density. The dashes (—) indicate a region in which no patch area is present. (Standardized to a total length of 0.5 mm.) Embryo 2 is shown in Fig. 2g, h.

number¹⁴, Held concluded that the positioning of the bristles in the adult cuticle is by a mechanism other than a positional gradient.

It appears that a different mechanism controls the spacing of the *ftz* stripes than is involved in the positioning of the bristles. If cell-cell interactions are involved in the establishment of the *ftz* pattern, these results make it clear that the pattern is not established through a cell counting mechanism. Even as a single stripe crosses a cellular density boundary in a mosaic embryo, the width of the stripe is not altered. Thus, the molecules involved in the expression of the *ftz* protein possess propagation properties which are independent of cell size and density.

These results support a model in which cortical cytoplasmic factors are involved in establishing the *ftz* striped pattern. Cytoplasmic cues readily explain the independence of the *ftz* expression with respect to cell size and density and its consistency with respect to the proportions of the entire embryo. They would not necessarily have to provide a detailed matrix delineating the *ftz* striped pattern, but only a very general regionalization; enough to establish only the proper proportions as the *ftz* stripes are being established and refined. In addition, these cues may not act directly on *ftz* gene expression, but indirectly through other segmentation loci which act before or in parallel with the *ftz* gene and influence its expression¹⁵.

I cannot rule out the alternative possibility that the pattern of *ftz* expression may already be established before nuclear cycle 14, even if it is not necessarily stably determined¹. Because there is little nuclear mixing, the pattern could be preserved in successive cell divisions¹⁶. In the mutants *dal* and *mh1182*, the pattern of *ftz* expression could therefore be established before the defects that result in abnormal densities are manifest. In this view, the molecules involved in establishing the *ftz* pattern could still operate by a nuclear counting mechanism but at an earlier time in development.

The embryos with increased cell density, derived from females homozygous for the *mh1182* allele, often produce cuticle patterns with a reduced or a normal number of segments¹⁷. Although this result is in accord with the normal *ftz* pattern found in haploid embryos, it is not necessarily expected. There exist a number of zygotic and maternal-effect mutations in which the resulting altered cuticle pattern is not accurately reflected by the *ftz* pattern at blastoderm^{15,18}. In addition, the relation between the *ftz* pattern and the cuticle pattern is complicated by the regulative abilities of the *Drosophila* embryo^{1,19,20}.

In any case, it is evident that the number, spacing and width of the *ftz* stripes are independent of cell size and density. In animals in general, haploid or polyploid individuals maintain

the same body and organ size although their cell size is either smaller or larger than that of the equivalent diploid cells²¹. Fankhauser's studies on polyploid salamander larvae demonstrated that the morphogenesis of normal structures occurs not only through a reduction in cell number but, even more striking, through alterations in cell shape²². It is clear that for many organisms the size of the units of development (including specific organs and the *fit* stripes) are not linked to cell size.

Received 8 December 1986; accepted 13 March 1987.

1. Simcox, A. A. & Sang, J. H. *Dev. Biol.* **97**, 212-221 (1983).
2. Scott, M. P. & O'Farrell, P. H. *A. Rev. Cell Biol.* **2**, 49-80 (1986).
3. Carroll, S. B. & Scott, M. P. *Cell* **43**, 47-57 (1985).
4. DiNardo, S., Kunter, J. M., Theis, J. & O'Farrell, P. H. *Cell* **43**, 59-69 (1986).
5. Gans, M., Audit, C. & Masson, M. *Genetics* **81**, 683-704 (1975).
6. Zalokar, M., Audit, C. & Erk, I. *Dev. Biol.* **47**, 419-432 (1973).
7. Sandler, L. *Genetics* **86**, 567-582 (1977).
8. Foe, V. & Alberts, B. J. *Cell Sci.* **61**, 31-70 (1983).
9. Edgar, B. A., Kiehle, C. P. & Schubiger, G. *Cell* **44**, 365-372 (1986).
10. Wang, J. *Molec. cell. Biol.* **5**, 3640-3643 (1985).
11. Held, L. I. *Willhelm Roux Arch. dev. Biol.* **187**, 105-127 (1979).

I thank M. Scott, S. Carroll and J. Wong for the antibodies used and T. Kornberg, P. O'Farrell, G. Schubiger, M. Kirschner and the members of their laboratories for their advice. I especially thank Bruce Alberts, in whose laboratory this work was carried out, for his advice and reading of the manuscript. I also thank E. Sullivan and Kathleen Rañeses for preparation of the manuscript. This work was supported by a Helen Hay Whitney Fellowship and by the National Institutes of Health.

12. Santamaria, P. *Dev. Biol.* **96**, 285-295 (1983).
13. Cavalier-Smith, T. *A. Rev. Biophys. Bioengng* **11**, 273-302 (1982).
14. Wolpert, L. *J. theor. Biol.* **25**, 1-47 (1969).
15. Carroll, S. B. & Scott, M. P. *Cell* **45**, 113-126 (1986).
16. Zalokar, M., Erk, I. & Santamaria, P. *Cell* **19**, 133-141 (1980).
17. Edgar, B. A., Kiehle, C. P. & Schubiger, G. *Cell* **44**, 365-372 (1986).
18. Carroll, S. B., Winslow, G. M., Schubach, T. & Scott, M. P. *Nature* **323**, 278-280 (1986).
19. Howard, K. & Ingham, P. *Cell* **44**, 949-957 (1986).
20. Gergen, J. P. & Wieschaus, E. F. *Dev. Biol.* **109**, 321-335 (1985).
21. Fankhauser, G. in *Analysis in Development* (eds Willier, B. H., Weiss, P. A. & Hamburger, V.) 126-150 (Hafner, New York, 1955).
22. Fankhauser, G. *J. exp. Zool.* **100**, 445-455 (1945).
23. Karr, T. & Alberts, B. M. *J. Cell Biol.* **102**, 1494-1509 (1986).

Ethane formation from acetylene as a potential test for vanadium nitrogenase *in vivo*

Michael J. Dilworth*, Robert R. Eady, Robert L. Robson & Richard W. Miller†

AFRC Unit of Nitrogen Fixation, University of Sussex, Brighton BN1 9RQ, UK

* School of Environmental and Life Sciences, Murdoch University, Murdoch, 6150 Western Australia

† Chemistry and Biology Research Institute, Agriculture Canada, Ottawa, K1A 0C6 Canada

Azotobacter chroococcum and *A. vinelandii*, both obligately aerobic diazotrophic bacteria, have at least two genetically distinct systems for nitrogen fixation^{1,2}. One is the well-characterized molybdenum nitrogenase, whereas in a second system (the vanadium nitrogenase) the conventional molybdoprotein is replaced by a vanadoprotein and a second iron protein replaces the one normally found with the molybdoprotein³⁻⁵. A third system may exist in *A. vinelandii*^{6,7}. Reiterations of DNA encoding nitrogenase structural genes (*nifHDK*) are found in several genera of diazotrophs (see ref. 1). In the azotobacters there is evidence, based on physiological and biochemical properties of strains carrying deletions of *nifHDK*, that the reiterated genes are involved in the synthesis of alternative nitrogenases including the vanadium (V)-nitrogenase. The screening of other genera for the presence of a functional V-nitrogenase would be greatly facilitated by a specific test for it which could be applied *in vivo*. The reduction of acetylene to ethylene^{8,9} by the molybdenum (Mo)-nitrogenase has been widely used in ecological, genetic and physiological studies of biological nitrogen fixation. We demonstrate here that purified V-nitrogenase catalyses the reduction of acetylene not only to ethylene, but also to ethane, which the Mo-nitrogenase does not¹⁰. This property distinguishes between these two systems in *A. chroococcum* and *A. vinelandii* *in vivo*. Ethane formation also occurred in cultures of *Clostridium pasteurianum* W5, suggesting that V-nitrogenase is not restricted to the *Azotobacteriaceae*.

Studies with strains of *A. chroococcum* and *A. vinelandii* deleted for the structural genes encoding the Mo-nitrogenase (*nifHDK*) have shown that during growth in Mo-deficient media, when the V-nitrogenase is produced, acetylene is a poor substrate for nitrogenase compared with H⁺ or N₂ (refs 1 and 11). The purified V-nitrogenase, but not the Mo-nitrogenase, shows this substrate reduction pattern³⁻⁵. This pattern of substrate reduction is also shown by wild-type *A. vinelandii* UW136,

but only when grown under molybdenum deficiency¹² where the Mo-nitrogenase does not function¹¹. A good correlation therefore exists between the molybdenum or vanadium nutritional status of the cultures and the characteristics of the nitrogenase present in them.

During an investigation of the substrate specificity of the V-nitrogenase of *A. chroococcum* we detected an additional gaseous product on gas chromatographic analysis of the head-space of acetylene reduction assays routinely used for ethylene analysis¹³. The product was identified as ethane by its retention time on an activated alumina column¹⁴ and by mass spectrometry. Ethane is a minor product accounting for no more than 3% of the electron flux through the V-nitrogenase.

Table 1 shows that ethane formation is a property of the VFe protein (Ac1*) because combinations of the Fe protein of the V-nitrogenase (Ac2*) with MoFe protein of the Mo-nitrogenase (Ac1) did not produce ethane, nor did homologous Mo-nitrogenase components of *A. chroococcum* or *Klebsiella pneumoniae*. Increasing the electron flux through the VFe protein of the V-nitrogenase by increasing the ratio of Ac2* to Ac1* results in an increased ethane/ethylene ratio (Table 1). This observation suggests that ethylene release is slower from the V-nitrogenase than from the Mo-nitrogenase, which would account for acetylene being a relatively poor substrate for the V-nitrogenase^{3,4}.

Ethylene itself was reduced to ethane by the V-nitrogenase, but only at about 40% of the rate of ethane formation from acetylene. Because commercial ethylene has such a high background of ethane that extensive purification is needed before ethane formation can be accurately measured, ethylene is not a suitable substrate for routine assay of V-nitrogenase activity. In contrast, acetylene generated from carbide has a negligible ethane background.

A. chroococcum (R.L.R., unpublished data) and *A. vinelandii*¹⁵ express the V-nitrogenase only under conditions of molybdenum deprivation because 50 nM Mo prevents synthesis of these enzymes. After three sub-cultures in medium containing no added Mo, cultures of *A. vinelandii*, *A. chroococcum* and *C. pasteurianum* formed ethane from acetylene (Table 2). Except in the case of *A. vinelandii* CA11, the addition of 1 µM VOSO₄ increased the proportion of ethane to ethylene by four- to fivefold. The low ethane/ethylene ratio observed in the unsupplemented medium can be explained by the simultaneous functioning of the Mo- and V-nitrogenases. However, in the case of CA11, the highest ratio was observed in unsupplemented medium. Because we have shown that the ethane/ethylene ratio can be affected by electron flux (Table 1), this may mean that the V-nitrogenase in CA11 is operating at very high flux. An alternative explanation is that if another nitrogenase is present in CA11 grown in unsupplemented medium, it must also reduce

Table 1 Association of ethane production from acetylene with the VFe protein of vanadium nitrogenase

Experiment	Protein combination	Ratio of component 2: component 1	Product nmol min ⁻¹			$\frac{C_2H_6}{C_2H_4} \times 100$
			C ₂ H ₄	C ₂ H ₆	H ₂	
1	Ac2 + Ac1	6.4	103.7	0	5.3	0
	Ac2* + Ac1	6.0	99.5	0	5.3	0
	Ac2 + Ac1*	5.2	38.3	0.75	41.2	2.0
	Ac2* + Ac1*	4.8	33.9	0.68	37.3	2.0
2	Ac2* + Ac1*	1.0	2.2	0.013	3.2	0.6
	Ac2* + Ac1*	20.0	33.4	0.74	48.4	2.2

V-nitrogenase component proteins (Ac1* and Ac2*) and Mo-nitrogenase components (Ac1 and Ac2) were purified as described in (refs 3 and 20). Nitrogenase activity was measured as described in ref. 13, except that C₂H₆, C₂H₄ and C₂H₂ were separated on an alumina column (1 m × 6 mm diameter) prepared as described in ref. 14 and operated at 110 °C with N₂ (45 ml min⁻¹) as the carrier gas²¹. The amounts of protein assayed were: experiment 1: Ac1, 105 µg; Ac2, 205 µg; Ac1* 112 µg; Ac2*, 165 µg. Experiment 2: Ac1* 65 µg and Ac2*, 18.4 or 368 µg.

Table 2 Ratio of ethane to ethylene produced by cultures of *A. vinelandii*, *A. chroococcum* and *C. pasteurianum* supplemented with V or Mo

	C ₂ H ₆ /C ₂ H ₄ (%)		
	Culture supplement		
	None	V	Mo
<i>Azotobacter chroococcum</i> MCD50	0.25	1.9	0
<i>Azotobacter chroococcum</i> MCD1155	no growth	2.5	no growth
<i>Azotobacter vinelandii</i> UW136	0.33	1.5	0.0002
<i>Azotobacter vinelandii</i> CA11	3.9	1.9	no growth
<i>Clostridium pasteurianum</i> W5	0.30	0.92	0
<i>Clostridium pasteurianum</i> ATCC6013	0.13	0.49	0

The azotobacter strains were *A. chroococcum* MCD50 (ref. 1), *A. vinelandii* UW136 (ref. 22), both wild-type for nitrogen fixation; *A. vinelandii* CA11 and *A. chroococcum* MCD1155 are strains deleted for the structural genes of Mo-nitrogenase ($\Delta nifHDK$) and unable to grow in the presence of Mo (refs 1 and 2). The clostridial strains were *C. pasteurianum* ATCC6013 obtained from NCIB and *C. pasteurianum* W5, nominally the same. Strains were grown on the nitrogen-free media of ref. 23, and the azotobacter strains on the media of ref. 3 supplemented with 100 µM Na₂ MoO₄ or 1 µM VOSO₄ as indicated, for three serial sub-cultures. Cultures were grown for 18 h and headspace gas analysed as described in Table 1 1–2 h after the injection of C₂H₂ to give 0.1 atm.

ranged from 4 to 6 rather than the 3 to 4 found for wild-type acetylene to ethane. The addition of 100 µM molybdate virtually eliminated ethane formation in all strains (Table 2). In the case of the *nifHDK* *A. chroococcum* MCD1155 deletion strain, molybdate also prevented growth. The two clones of *C. pasteurianum*, obtained from different sources, show a clear difference in the ethane/ethylene ratio. The reasons for this difference are not clear, but may reside in the maintenance history of the clones. In both cases, however, molybdate completely prevented ethane production from acetylene.

Tests of various acetylene concentrations showed that 0.2 atm was saturating for all three organisms, and gave maximal ethane/ethylene ratios. The nitrogenase activity of the azotobacter strains is strongly influenced by dissolved oxygen, but the ethane/ethylene ratio did not vary with the partial pressure of oxygen (pO_2) (data not shown).

Despite the low allocation of electrons to ethane formation, this reaction provides a satisfactory assay for the V-nitrogenase *in vivo*. Ethane elutes first during chromatography, is well resolved from ethylene and acetylene, and is readily detectable at 40 pmol per injection. As an assay, it is sensitive, uncomplicated by pO_2 effects in the case of aerobes and uses an easily accessible concentration of acetylene already widely used in physiological studies on biological nitrogen fixation. It has given a positive response with the two organisms known to form the V-nitrogenase (*A. chroococcum* and *A. vinelandii*). Molybdenum inhibition of ethane formation by *C. pasteurianum* is indicative of a V-nitrogenase in this species. This is consistent with an early report of stimulation by vanadium of N₂ fixation by *Clostridium* species¹⁶, implying that one of the reiterated *nifH* genes in *C. pasteurianum*¹⁷ codes for the Fe protein of a V-nitrogenase.

Before this technique can be used with confidence with organisms other than azotobacter strains, more examples are needed of a positive correlation between ethane formation and V supply in Mo-deficient conditions. The ethane assay clearly differentiates between the V- and Mo-nitrogenases, but the possibility remains that other nitrogenases such as the putative third nitrogenase of *A. vinelandii*⁶, may prove to have this property.

We hope that this report will stimulate appropriate tests for alternative nitrogenases in other diazotrophic organisms, particularly those possessing reiterations of structural genes for nitrogenase, for example *Rhizobium*¹⁸ and *Rhodobacter*¹⁹.

We thank Professor J. R. Postgate for criticism of the manuscript, Mrs M. E. Eldridge for technical assistance and Dr P. Bishop for strain CA11.

Received 13 February; accepted 18 March 1987.

- Robson, R. L. *Arch. Microbiol.* **146**, 74–79 (1986).
- Bishop, P. E. *et al. Science* **232**, 92–94 (1986).
- Robson, R. L. *et al. Nature* **322**, 388–390 (1986).
- Hales, B. J., Case, E. E., Morningstar, J. E., Dzieda, M. F. & Mauteser, L. A. *Biochemistry* **25**, 7251–7255 (1986).
- Hales, B. J., Langosch, D. & Case, E. E. *J. biol. Chem.* **261**, 15301–15306 (1986).
- Joerger, R. D., Premakumar, R. & Bishop, P. E. *J. Bact.* **168**, 673–682 (1986).
- Bishop, P. E., Jarlenski, D. M. L. & Hetherington, D. R. *J. Bact.* **150**, 1244–1251 (1982).
- Dilworth, M. J. *Biochim. biophys. Acta* **127**, 285–294 (1966).
- Schöllhorn, R. & Burris, R. H. *Proc. natn. Acad. Sci. U.S.A.* **58**, 213–216 (1967).
- Hardy, R. W. F., Holsten, R. D., Jackson, E. K. & Burns, R. C. *Pl. Physiol.* **43**, 1185–1207 (1968).

- Bishop, P. E., Hawkins, M. E. & Eady, R. R. *Biochem. J.* **238**, 437–442 (1986).
- Eady, R. R. & Robson, R. L. *Biochem. J.* **224**, 853–862 (1984).
- Eady, R. R. *Meth. Enzym.* **69**, 751–776 (1980).
- Smith, K. A. & Dowdell, R. J. *J. chromat. Sci.* **11**, 655–658 (1973).
- Jacobson, M. R., Premakumar, R. & Bishop, P. E. *J. Bact.* **167**, 480–486 (1986).
- Jensen, H. L. & Spencer, D. *Proc. Linn. Soc. N.S.W.* **72**, 73–86 (1947).
- Chen, K.-C.-K., Chen, J.-S. & Johnson, J. L. *J. Bact.* **166**, 162–172 (1986).
- Quinto, C. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 1170–1174 (1985).
- Scolnik, P. A. & Haselkorn, R. *Nature* **307**, 289–292 (1984).
- Yates, M. G. & Planque, K. *Eur. J. Biochem.* **60**, 467–476 (1975).
- Smith, K. A. & Restall, S. W. *J. Soil Sci.* **22**, 430–443 (1971).
- Bishop, P. E., Gordon, J. K., Shah, V. K. & Brill, W. J. in *Genetic Engineering for Nitrogen Fixation* (ed. Hollaender, A.) 67–80 (Plenum, New York, 1977).
- Winogradsky, S. *Arch. sci. biol.* **3**, 297–352 (1895).

Inequality in mutation rates of the two strands of DNA

Chung-I Wu* & Nobuyo Maeda†

* Department of Biology, University of Rochester, Rochester, New York 14627, USA

† Laboratory of Genetics, University of Wisconsin, Madison, Wisconsin 53706, USA

As the mechanisms for replicating the two strands of duplex DNA differ¹ it is, in principle, possible for the mutation rates to differ depending on which strand is being copied. In the absence of selection this would lead to a difference in the measured rate of a particular base substitution, such as T to C, depending on which DNA strand was analysed to determine the rate. Thus a change such as T to C on one DNA strand results from either a direct T-to-C mutation on that strand or an A-to-G mutation on the complementary strand; for the other strand the situation is reversed, and it can be seen that different processes are responsible for the two cases, allowing for asymmetry in substitution rate. We have tested whether such asymmetry indeed occurs by studying equivalent sequences from the β -globin complexes of six species of primate. Our results reveal an asymmetry in substitution rates consistent with predictions based on strand-inequalities in mutation rates. Our sequence comparisons also allow us to make predictions about the positions of replication origins and the replication error rates of one strand relative to the other.

Let us designate the two strands of DNA as S (for sense strand) and AS (for antisense strand); the designation is, of course, independent of the manner in which each is replicated. Nucleotide substitutions are scored on the S strand. A change from T→C actually means a T·A pair mutates to a C·G pair. The mutation rate of T→C, designated M_{TC} , is therefore the sum of rates of replication errors of T→C on the S strand and A→G on the AS strand. Similarly, M_{AG} is the sum of rates of A→G on the S strand and T→C on the AS strand. If the patterns and rates of replication error of the two strands are identical, we expect $M_{TC} = M_{AG}$; similarly $M_{CT} = M_{GA}$, $M_{CA} = M_{GT}$, $M_{AC} = M_{TG}$, $M_{GC} = M_{CG}$ and $M_{AT} = M_{TA}$. Such symmetries are not affected by the methylation of the C of the 5'CG3' dinucleotide which results in an elevated rate of C→T or G→A (ref. 2). (An alternative way to consider this question is to imagine the possible substitution patterns if one strand never mutates.)

To study strand asymmetry, it is necessary to analyse a contiguous stretch of DNA from several species whose phylogenetic relationships are unambiguous. These sequences should also be small enough to be within a single replication unit. We analysed seven sequences from six species of primates: human (*Homo sapiens*, 2 sequences), chimpanzee (*Pan troglodytes*), gorilla (*Gorilla gorilla*), orang utan (*Pongo pygmaeus*), rhesus monkey (*Macaca mulatta*, an Old World monkey) and spider monkey (*Ateles geoffroyi*, a New World monkey). The sequence is a 3.1-kilobase (kb) fragment between the $\psi\beta$ -globin (or $\psi\eta$ -globin (ref. 3)) and the δ -globin genes of the β -like haemoglobin gene complex^{4,5}. This region is non-repetitive except for two *Alu* inserts and a region very rich in T and C (on one strand). The region has no known function and is apparently only mildly constrained by selection⁵. The phylogenetic relationships of the seven sequences can be unambiguously determined⁵. The ancestral sequences and the types of substitutions that occurred can be inferred by using standard procedures^{6,7}. In total, there are 507 substitutions scored. In Table 1, we give the relative rates (expressed as percentages), f_{IJ} , for all twelve types of substitutions (I→J, where I, J each stand for one of the bases T, C, A or G and I≠J). The relative rate has been adjusted for the unequal representation of each nucleotide in the sequences and is the percentage for each of the twelve types of nucleotide substitutions among all substitutions⁷. Table 1 confirms some

Table 1 Percentage for each of the 12 types of nucleotide substitutions among all substitutions

All substitutions		J				
		T	C	A	G	Sum
I	T	—	14.1	2.4	3.2	19.7
	C	21.1	—	4.9	2.4	28.5
	A	3.5	4.0	—	14.1	21.6
	G	5.8	6.5	17.7	—	30.0
	Sum	30.4	24.7	25.0	19.8	100.0
Transversions		T or C		A or G		—
	T or C	—		12.9 (12.3)		
I	A or G	19.8 (18.9)				

The direction of changes is from the nucleotide I (column) to the nucleotide J (row). Summarized in the lower part of the table, changes from a purine to a pyrimidine are in bold type and changes from a pyrimidine to a purine are in italic. The numbers in parentheses are obtained from a subset of the data that includes only changes leading to the Old World monkeys and the apes.

previous observations, including higher rates for transitions, and, among transitions, higher rates for (C·G)→(T·A) than (T·A)→(C·G)⁸.

To examine the mutation rates of the two strands, we may consider the two ratios for transitions, $f_{TC}/f_{AG} = 14.1/14.1 = 1.0$ and $f_{CT}/f_{GA} = 21.1/17.7 = 1.19$, and the ratio for transversions $f_{RY}/f_{YR} = 19.8/12.9 = 1.53$. Here, R represents a purine and Y a pyrimidine so that $f_{RY} = (f_{AT} + f_{AC} + f_{GT} + f_{GC})$ and $f_{YR} = (f_{TA} + f_{TG} + f_{CA} + f_{CG})$. Neither of the two transition ratios is significantly different from 1 whereas R→Y occurs significantly more frequently than Y→R. There are 81.5 substitutions of R→Y, out of 1,314 purines (averaged over the seven sequences), and 77.5 substitutions of Y→R, out of 1,840 pyrimidines. In a two-tailed binomial test, the difference is significant ($P = 0.013$). In the comparisons of individual pairs, changes from a purine (A or G) to a pyrimidine (T or C) have been consistently more frequent than the corresponding changes from the pyrimidine to the purine. In one particular comparison (f_{GC} versus f_{CG}), one is three times higher than the other. The differences, 11.8 C→G changes out of 750 C residues and 20.2 G→C changes out of 476 G, are highly significant ($P = 0.002$). In the $\psi\beta$ -globin region about 4 kb upstream analysed previously by Koop *et al.*³, we can see the same pattern. In their analysis f_{GC} (5.9%) and f_{CG} (3.1%) is also the pair most deviant from equality and f_{RY}/f_{YR} is again more deviant from equality than either f_{TC}/f_{AG} or f_{CT}/f_{GA} .

Asymmetry in mutation rate and/or mutation pattern between the two strands should be reflected in nucleotide compositions as well. If symmetry holds true, we expect the equilibrium frequencies of nucleotides to be the same for both strands. In other words, the frequency of T should equal the frequency of A on the same strand. Similarly, the frequency of C and G should be the same. The nucleotide composition of the S strand is given in Table 2. We can see that the symmetry does not hold.

Table 2 Nucleotide composition of the S strand

	T	A	C	G
Entire sequences	34.6% > 26.6%	23.8% > 15.1%		
Unique parts of sequences*	35.3% > 29.4%	21.4% > 13.9%		
Prediction of Table 1	35.5% > 23.2%	24.7% > 16.7%		

* Excludes the two *Alu* inserts which are rich in G and C and the region of the last 366 bp which is rich in T and C.

Table 3 Net changes of nucleotides in different lineages

Lineage	ΔT	ΔA	ΔC	ΔG
Old World monkeys	+6 >	-13	-4 <	+11
Apes	-16 <	-10	+16 >	+10

Values are the net changes of nucleotides along each branch. For example, along the Old World monkey branch, there have been 26 changes from T and 32 changes to T since the time of divergence from the apes, resulting in a net change of +6. Bold signs mark in equalities in the same direction as the imbalance in nucleotide composition (see Table 2), indicating further departure from the symmetric composition with elapsed time.

In addition, scoring for the gain and loss of each nucleotide along each branch of the phylogenetic tree reveals no trend towards symmetry either (Table 3). Table 2 also gives the equilibrium nucleotide frequencies as determined⁹ by the mutation matrix of Table 1. The observed composition differs somewhat from the expected frequencies; however, the observed inequalities between T and A and between C and G are as expected ($T > A$ and $C > G$). The derivation of the substitution pattern of Table 1 is independent of the nucleotide composition, so this agreement may not be a coincidence.

Selection is not likely to account for the observed asymmetry. If selection has led to a bias against $Y \rightarrow R$ changes, the total percentage of transversions in this region (33%) should have been lower than that in the nearby $\psi\beta$ -globin region (32.6%). The overall level of selection in this region is also too weak to upset symmetry by as much as observed: the total divergence among the apes and the two monkeys is 16.0% in comparison with 17.5% for the $\psi\beta$ -globin region³. If selection acted strongly against $Y \rightarrow R$ changes to produce the observed asymmetry, we should have expected the level of divergence to be lower than that observed. In addition, the overall similarity of the asymmetric substitution pattern in the $\psi\beta$ -genes³, which should have been neutral, argues against the influence of selection.

An adequate model must account for the finding that the transversion ratio is more deviant than either of the two transition ratios. Such a model can best be developed in the light of molecular mechanisms of nucleotide substitutions¹⁰. It has been shown experimentally that, on a given strand, $Y \rightarrow R$ mutations occur at a much lower frequency than other kinds of changes¹¹. The low frequencies of $Y \rightarrow R$ changes may be partially explained by the lack of suitable tautomers for $Y \cdot Y$ mismatches¹⁰. We may state the transversion ratio, f_{RY}/f_{YR} , and the two transition ratios quantitatively. Let α_{IJ} be the rate of mutation from nucleotide I to J on the S strand and β_{IJ} be the rate on the AS strand. For simplicity, we shall let $\beta_{IJ} = k\alpha_{IJ}$ for all I and J, where k is a non-negative constant. For transversions, $f_{RY}/f_{YR} = (\alpha_{RY} + \beta_{YR})/(\alpha_{YR} + \beta_{RY}) \approx \alpha_{RY}/\beta_{RY} = 1/k$. For transitions, $f_{TC}/f_{AG} = (1 + kP)/(k + P)$, where $P = \alpha_{AG}/\alpha_{TC}$ and $f_{CT}/f_{GA} = (1 + kQ)/(k + Q)$, where $Q = \alpha_{GA}/\alpha_{CT}$. It is therefore clear that the ratio f_{RY}/f_{YR} should always deviate from 1 more than f_{TC}/f_{AG} or f_{CT}/f_{GA} , if the two strands have unequal replication fidelity. In addition, $f_{RY}/f_{YR} = 1/k$ should reflect the ratio of the replication errors on the S strand to the replication errors on the AS strand. This interpretation assumes a fixed origin of replication in the germ cells.

In the 3.1-kb region analysed, $f_{RY}/f_{YR} = 1/k = 1.53$. Because $k = 0.65$, we may conclude that the AS strand has a smaller replication error than the S strand. We hypothesize that the mutation asymmetry could be due to the replication characteristics of this region. In other words, our observations and those of a previous study³ suggest that a contiguous DNA region from the $\psi\beta$ gene to the 3' end of this 3.1-kb fragment may have been either consistently or at least predominantly on the same side of a replication origin (or origins). If the replication error of a leading strand is smaller than a lagging strand, we may predict

that the preferred replication origin is on the 5' side of this region. Furthermore, if the sequences from another region of β -globin gene complex are available, and show a substitution pattern opposite to that of Table 1, it would suggest the existence of a replication origin in between.

We thank James F. Crow for inspiration and encouragement, Uzi Nur for incisive criticisms, and Oliver Smithies, Marty Gorovsky and Ed Stephenson for helpful comments.

Received 2 November 1986; accepted 27 February 1987.

1. Lewin, B. *Genes* (Wiley, New York, 1985).
2. Razin, A. & Riggs, A. D. *Science* **210**, 604-610 (1980).
3. Koop, B. F., Goodman, M., Xu, P., Chan, K. & Slightom, J. L. *Nature* **319**, 234-238 (1986).
4. Maeda, N., Bliska, J. B. & Smithies, O. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5012-5016 (1983).
5. Maeda, N., Wu, C. I., Bliska, J. B. & Reneker, J. *Molec. Biol. Evol.* (submitted).
6. Fitch, W. M. *Syst. Zool.* **20**, 406-416 (1971).
7. Gojobori, T., Li, W. H. & Graur, D. *J. molec. Evol.* **18**, 360-369 (1982).
8. Lu, W. H., Wu, C. I. & Luo, C. C. *J. molec. Evol.* **21**, 58-71 (1984).
9. Tajima, F. & Nei, M. *J. molec. Evol.* **18**, 115-120 (1982).
10. Topal, M. D. & Fresco, J. R. *Nature* **263**, 285-293 (1976).
11. Fersht, A. R. & Knill-Jones, J. W. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4251-4255 (1981).

Release of mature starfish oocytes from interphase arrest by microinjection of human centrosomes

A. Picard*, E. Karsenti†, M. C. Dabauvalle† & M. Dorée†‡

* Laboratoire Arago, F-66650 Banyuls, France

† European Molecular Biology Laboratory, Postfach 10.2209, D-6900 Heidelberg, FRG

Mature oocytes (unfertilized eggs) are arrested at definite cell-cycle stages which vary from species to species. In frogs and mammals, the oocytes are arrested at the second metaphase of meiosis whereas in echinoderms they are blocked later, at the pronucleus stage. What causes the maturing oocytes to stop at some point in the cell cycle is not entirely clear. In frogs, the metaphase arrest seems to be maintained by a cytotstatic factor¹. In echinoderms, which stop at interphase, no such a factor has so far been found. The fertilization process, beyond the introduction of paternal chromosomes, releases the oocyte from cell-cycle arrest and provides a functional centrosome to replace the endogenous centrosome which is apparently lost during oogenesis in most species³⁻⁵. Several lines of evidence suggest that release from cell-cycle arrest is mediated by a Ca^{2+} burst which is associated with fertilization⁶⁻⁹, and it is known that the functional centrosome provided by the sperm is necessary for mitotic spindle formation and cleavages^{4,6,10}. We report here that microinjection of purified human centrosomes into mature starfish oocytes is sufficient to release them from arrest at interphase and to support many cleavages leading to the occasional formation of normal embryos. In this species centrosome induced re-entry into the cell cycle does not require a transient calcium burst nor does it require intact microtubules.

In starfish, as in many species, the oocyte is arrested in late G_2 . Meiotic maturation is triggered by 1-methyladenine, a hormone produced by the follicle cells¹¹. Fully grown oocytes treated with 1-methyladenine eliminate two polar bodies and arrest at interphase with the female pronucleus intact, containing a single set of decondensed chromosomes (Fig. 1a). At this stage, they lack a functional centrosome. Human centrosomes have been prepared from KE37 lymphoid cells according to a modification of the procedure described by Mitchison and Kirschner¹². The use of lymphoid cells and the addition of 0.5 mM

‡ Permanent address: CNRS, BP 5051, F-34033 Montpellier Cedex, France.

Fig. 1 Micrographs ($\times 190$) showing the location of condensed chromatin (as revealed with Hoechst dye 33258). *a*, Chromatin in a fully mature starfish oocyte arrested after formation of the female pronucleus. Arrows point to condensed chromatin, observed in each polar body, but not in the oocyte. *b*, The same in a cleaving embryo originating from a fully mature oocyte injected with 1–5 human centrosomes. The embryo was fixed 8 h after centrosome microinjection. *c*, The same in a fully mature oocyte injected with 1–5 centrosomes 20 min after transfer to sea water containing $10 \mu\text{g ml}^{-1}$ nocodazole. The cell was fixed 6 h after centrosome injection. Several chromosome cycles occurred without cleavage, producing a cell with a large number of chromosomes.

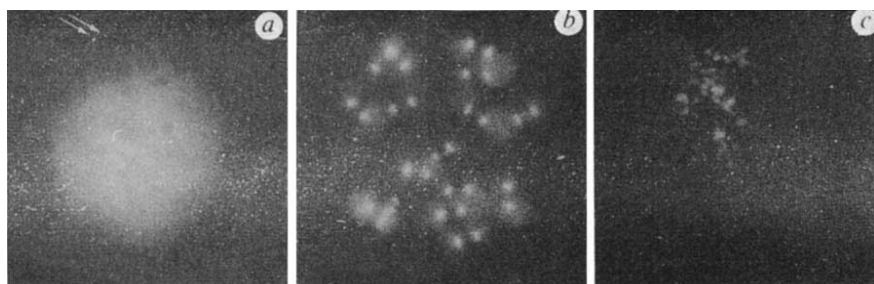
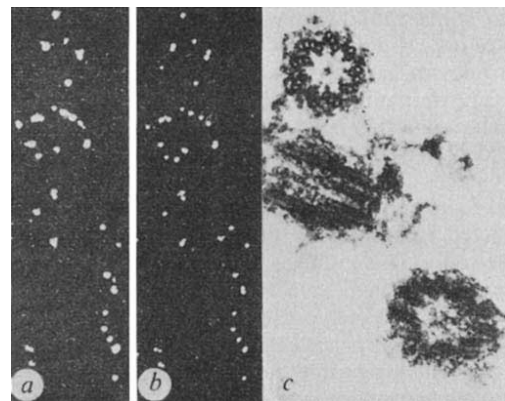


Fig. 2 Centrosomes isolated from human KE37 lymphoid cells. Centrosomes were centrifuged on to glass coverslips and stained by double immunofluorescence for pericentriolar material using the human anti-centrosome serum 5051 (ref. 12) (*a*) and a monoclonal antibody against β -tubulin (Amersham) (*b*). Centrosomes were counted according to Mitchison and Kirschner¹² on such preparations ($\times 1,000$). An electron micrograph ($\times 42,000$) of the same preparation is shown in part *c*.



MgCl_2 to the lysis buffer has led to a dramatic improvement of the purity of the centrosomes which are now close to homogeneity^{13,14}. The centrosomes are fractionated on a discontinuous sucrose gradient made up in a buffer containing 10 mM potassium PIPES pH 7.2, 1 mM EDTA, 0.1% β -mercaptoethanol and 0.1% Triton X-100. They peak in 50–55% sucrose at the interface between the 70% and 50% layers at about 10^8 centrosomes ml^{-1} . When mature *Marthasterias glacialis* starfish oocytes were injected with 50 picolitres of the undiluted centrosome fraction (1–5 centrosomes) they cleaved (Fig. 1*b*), normal embryogenesis ensued and gastrula were occasionally observed. This represents a small amount of centrosomes per egg and the frequency of cleavage observed (85%, 5 h after injection) (Table 1) is reasonable compared to that observed in frog eggs injected with similar amounts of centrosomes^{6,14}. The injection of smaller amounts of centrosomes resulted in a lower percentage of cleaving eggs. Injection of 5 pl of the centrosome fraction (less than one centrosome per egg) resulted in almost no cleavages. The eggs remained arrested at the pronucleus stage, as did control oocytes injected with a 50% sucrose solution prepared in the gradient buffer. These results show that contrary to what happens in frogs, the pricking event has no effect on mature starfish oocytes and strongly suggest that the centrosomes are necessary to release the oocytes from cell-cycle arrest. In a further effort to determine if the centrosomes were indeed the active component, the preparation was diluted 1:1 with 10 mM potassium PIPES pH 7.2 to reduce the sucrose concentration to 25% and centrifuged for 15 min at either 2,000g or 160,000g. Centrosomes do not sediment at 2,000g but do so at 160,000g. Aliquots (50 pl) of each supernatant were then injected into 10 oocytes. The 20 microinjected oocytes were fixed 2 h later and examined with the Hoechst dye 33258. Six out of the 10 oocytes injected with the 2,000g supernatant entered the cell cycle, as evidenced by the disappearance of the pronucleus and condensation of the chromosomes. The 10 oocytes injected with the 160,000g supernatant remained arrested at the pronucleus stage. These results

show that the activity which releases oocytes from cell-cycle arrest consistently cofractionates with the centrosomes.

Because simple pricking of starfish oocytes did not result in release from cell-cycle arrest, we tested the effect of the calcium ionophore A 23187, a potent universal parthenogenetic agent¹⁵. All the mature oocytes treated with this ionophore underwent several cycles of chromosome duplication and condensation without cleavage, producing eggs with large numbers of chromosomes. The same result was observed when oocytes were injected with 10 pl of a mixture of 0.5 M Ca-EGTA and 0.5 M EGTA (mixing ratio: CaEGTA/EGTA = 10) a calcium buffer which was calculated to contain $8 \mu\text{M}$ free Ca^{2+} (refs 16 and 17). These results show that in starfish, as in all other species, a calcium burst can release the mature oocyte from cell-cycle arrest. Moreover, it confirms that exogenous centrosomes are required for cleavage to occur. Pricking the oocyte does not release the cell-cycle arrest simply because, in starfish, calcium release is not induced by this method.

We have also tested whether injected centrosomes release cell-cycle arrest by producing a calcium burst when injected into mature oocytes. During the normal fertilization process as well as during activation induced by the ionophore A23187, the calcium wave is associated in starfish as well as in other species with cortical granule exocytosis and elevation of a fertilization membrane^{7,8,17,18}. This last phenomenon is easy to observe in the microscope and occurs only if the calcium burst has taken place. Neither centrosome buffer nor centrosome injection produced an elevation of the fertilization membrane although in the latter case the eggs entered the cell cycle and cleaved. Moreover, mature oocytes injected with centrosomes in the presence of EGTA also entered the cell cycle and cleaved (Table 1). On the other hand, EGTA microinjection completely abolished oocyte activation either by A23187 or by the microinjection of the same volume of the above calcium buffer (0.5 M; CaEGTA/EGTA ratio of 10:1). In eggs injected with centrosomes and EGTA, chromosome cycling and cleavage were

Table 1 Effect of a previous EGTA microinjection on induction of cell cycling by centrosome microinjection into mature oocytes

Treatment	Cleavage (%)	
	After 3 h	After 5 h
-EGTA	70 (14/20)	85 (17/20)
+EGTA	20 (5/25)	65 (13/20)

Mature oocytes were injected with 10 pl of a 0.5 M EGTA solution, or not injected. Then they were injected with 50 pl of a suspension containing 1–5 centrosomes. Five hours after the time of centrosome microinjection, a small proportion (<5%) of control uninjected oocytes had spontaneously escaped cell-cycle arrest. In the remaining oocytes (>95% of controls) an intact pronucleus containing no condensed chromatin was observed.

delayed to some extent compared to oocytes injected with centrosomes only (Table 1). We therefore considered the possibility that oocytes must first escape from the low pCa clamp (due to EGTA microinjection) before entering the mitotic cell cycle. To eliminate this possibility, mature oocytes injected or not injected with 10 pl of 0.5 M EGTA were further injected with centrosomes and treated at the two-cell stage with A 23187. Elevation of the fertilization membrane did not occur in the EGTA injected eggs, whereas it did so in the eggs injected with centrosomes only. This indicates that the embryos injected with EGTA are still clamped at a low pCa after the first mitotic cleavage. These experiments strongly suggest that centrosomes do not release the interphase cell cycle arrest through a calcium burst. It remained possible however, that the Ca^{2+} pulse necessary to release cortical granules was larger than that needed to start the cell cycle and that EGTA buffers only against high levels of Ca^{2+} . To test this possibility, we injected Ca^{2+} -EGTA buffers (10 pl, 0.5 M) of various CaEGTA/EGTA ratios to see if the level of free Ca^{2+} required to release oocytes from interphase cell cycle arrest was lower than that required to induce cortical granules exocytosis. Cortical granule exocytosis was observed in 50% of the eggs injected with a 5:1 Ca EGTA/EGTA buffer (20 injected eggs) whereas release from cell cycle arrest occurred in 50% of the eggs injected with a 7:1 CaEGTA/EGTA buffer (20 injected eggs). Therefore, release from cell cycle requires a higher free calcium raise than cortical granule exocytosis. Because the injection of EGTA blocks cell-cycle initiation by the calcium ionophore but not by centrosomes, we conclude that centrosomes release starfish mature oocytes from cell cycle arrest by a mechanism which does not involve calcium.

The most obvious function of centrosomes is to nucleate microtubules^{5,19}. It therefore seemed possible that cell-cycle arrest was released through the formation of microtubule asters nucleated by the centrosomes injected in the interphase cyto-

plasm of the mature oocytes. To test this possibility, we first investigated the effect of a potent microtubule inhibitor, nocodazole, on centriole reinitiation of the cell cycle. This drug was found to suppress cleavage in 50% of the fertilized eggs at 3 ng ml⁻¹ and in all fertilized eggs at 5 ng ml⁻¹. Therefore, starfish eggs are at least as sensitive to nocodazole as sea urchin eggs, where the drug has been shown to suppress microtubule polymerization at 20 ng ml⁻¹ (ref. 20). We incubated unfertilized starfish eggs in sea water containing 10 µg ml⁻¹ nocodazole for 20 min. They were then microinjected with 50 pl of human centrosomes. Pronuclear envelope disappeared and chromosomes condensed with the same frequency (80%, 20 eggs injected) as in eggs injected with centrosomes in the absence of nocodazole. The oocytes entered successive cycles of chromosome replication and condensation but did not cleave because the drug suppressed cytokinesis (Fig. 1c). This strongly suggests that, in starfish oocytes, centrosome induced release of cell-cycle arrest is not mediated by a microtubule-dependent process.

The mechanism by which the cell cycle is blocked at specific stages after meiosis reinitiation is completely unclear and very little studied in molecular terms. In frog eggs, the metaphase block maintained by the still uncharacterized cytostatic factor¹ is definitely released by the calcium wave which results from fertilization or pricking⁶. In this case, only the metaphase arrested eggs can be fertilized. In starfish, fertilization normally takes place while the oocyte is going through the meiotic cycles. In the absence of the sperm, the oocyte does not arrest at metaphase II but at the pronucleus stage (interphase). This interphase block may not be positively maintained but rather due to the lack of an active factor which is required for further progression in the cell cycle. The centrosomes we injected were isolated from non-synchronized cultured lymphoid cells. They may contain this component in an active form. The release from cell-cycle arrest which is induced by the calcium ionophore could be mediated by a calcium-dependent event which activates a stored centrosome component. Several reports support the view that centrosomes coordinate nuclear and cytoplasmic divisions by providing a signal for cell cleavage, which requires microtubule integrity^{21–23}. We confirm this view and suggest that, independently of their microtubule organizing function, centrosomes might provide a signal which is necessary for progression from the meiotic to the mitotic cell cycle in starfish. A considerable body of evidence supports the view that transition from interphase to mitosis depends on protein phosphorylation^{24–27}. It is perhaps relevant to this study that a monoclonal antibody directed against mitotic phosphoproteins has been shown to detect centrosomes²⁸ and that the catalytic and regulatory subunits of cyclic AMP dependent protein kinase are associated with centrosomes and the Golgi complex^{29,30}.

We thank K. Simons, B. Burke, J. Davoust and T. Kreis for critical reading of the manuscript, and Annie Steiner for typing.

Received 16 February; accepted 12 March 1987.

1. Meyerhof, P. G. & Masui, Y. *Dev Biol.* **72**, 182–187 (1979).
2. Szollosi, D., Calarco, P. & Donahue, P. J. *Cell Sci.* **11**, 521–541 (1972).
3. Callarco-Gillam, P., Siebert, M., Hubble, R., Mitchison, T. & Kirchner, M. W. *Cell* **35**, 621–629 (1983).
4. Schatten, H. & Schatten, G. *Cell Motility* **6**, 163–175 (1986).
5. Karsenti, E. & Maro, B. *Trends biochem. Sci.* **11**, 460–463 (1986).
6. Karsenti, E., Newport, J., Hubble, R. & Kirchner, M. J. *Cell Biol.* **98**, 1730–1745 (1984).
7. Zucker, R. S. & Steinhardt, R. A. *Biochem biophys. Acta* **541**, 459–466 (1978).
8. Ridgeway, E. B., Gilkey, J. C. & Jaffe, L. F. *Proc. natn. Acad. Sci. U.S.A.* **74**, 623–627 (1977).
9. Roy Cuthbertson, K. S., Whittingham, D. G. & Cobbold, P. H. *Nature* **294**, 754–757 (1981).
10. Kuriyama, R. & Borisy, G. G. *J. Cell Biol.* **91**, 814–821 (1981).
11. Kanatani, H., Shirai, H., Nakanishi, K. & Kurokawa, T. *Nature* **211**, 273–277 (1969).
12. Mitchison, T. & Kirchner, M. W. *Nature* **312**, 232–237 (1984).
13. Gosti-Testu, F., Marty, M. C., Berges, J., Maunoury, R. & Bornens, M. *EMBO J.* **5**, 2545–2550 (1986).
14. Bornens, M., Paintrand, M., Berges, J., Marty, M. C. & Karsenti, E. *Cell Motility* (in the press).

15. Steinhardt, R. A., Epel, D., Carroll, E. J. & Yanagimachi, R. *Nature* **252**, 41–43 (1974).
16. Portzehl, H., Caldwell, P. C. & Rüegg, J. C. *Biophys. biochim. Acta* **79**, 581–591 (1964).
17. Hamaguchi, Y. & Hiramoto, Y. *Expl Cell Res* **134**, 171–180 (1981).
18. Vacquier, V. D. *Dev Biol.* **84**, 1–26 (1981).
19. Kirchner, M. & Mitchison, T. *Cell* **45**, 329–342 (1986).
20. Schatten, G. & Schatten, H. *Expl Cell Res* **135**, 311–330 (1981).
21. Mitchison, J. M. *Symp. Soc. exp. Biol.* **6**, 105–125 (1952).
22. Hiramoto, Y. *Expl Cell Res* **68**, 291–298 (1971).
23. Swann, M. M. & Mitchinson, J. M. *J. exp. Biol.* **30**, 506–514 (1953).
24. Miake-Lye, R. & Kirchner, M. W. *Cell* **41**, 165–175 (1983).
25. Dorée, M., Peaucellier, G. & Picard, A. *Dev. Biol.* **99**, 489–501 (1983).
26. Burke, B. & Gerace, L. *Cell* **44**, 639–652 (1986).
27. Simanis, V. & Nurse, P. *Cell* **45**, 261–268 (1986).
28. Vandre, D. D., Davis, F. M., Rao, P. N. & Borisy, G. G. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4439–4443 (1984).
29. Nigg, E. A., Schäfer, G., Hilz, H. & Eppienberger, H. H. *Cell* **41**, 1039–1051 (1985).
30. De Camilli, P., Moretti, M., Domini, S., Walter, U. & Lohman, S. M. *J. Cell Biol.* **103**, 189–203 (1986).